Establishing murine embryonic stem cells for generation of transgenic mice featuring predictable and regulatable transgene expression

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades einer

Doktorin der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Pamela Riemer
aus Uelzen

1. Referent: Prof. Dr. Jürgen Wehland
2. Referent: Prof. Dr. Michael Steinert
eingereicht am: 22.10.2007
mündliche Prüfung (Disputation) am: 28.02.2008

Druckjahr 2008

Summary

Transgenic mice provide means to dissect and model events during developmental processes or human diseases. They are commonly generated by Random Integration (RI) of the transgene into the genome of murine Embryonic Stem (mES) cells and subsequent injection or aggregation of manipulated mES cell clones into/with early mouse embryos. The chromosomal integration site of the gene of interest has a strong impact on expression patterns in the animal ("position effect"), which is in particular critical for regulated transgene expression. Hence, generation of a transgenic mouse model is hindered by high screening efforts to identify an integration site supporting adequate (and well regulated) expression levels in all tissues required. Strategies devised to circumvent the problems of RI, e.g. site specific integration by homologous recombination or integration of large Bacterial Artifical Chromosomes, suffer from various limitations like the limited number of suitable integration sites currently available and difficult handling.

Consequently, the aim of this work was to screen for chromosomal loci in mES cells that allow for high and well regulated transgene expression in mice and to render them reusable via Recombinase Mediated Cassette Exchange (RMCE). The screening cassette applied for this purpose contained an autoregulatory positive feedback loop driven by the reverse transactivator of the tet system (rtTA). Since the properties of autoregulated transgene expression have not yet been examined in detail on transgenic mouse level, evaluation of this mode of regulated expression was a major issue of this work.

For targeting chromosomal loci previously tagged with Flp recombinase recognition target sequences via RMCE, a puromycin acetyl transferase/Δneomycin phosphotransferase selection strategy could be shown to work highly efficient in mES cells.

Random tagging of sites with the screening cassette yielded low numbers of clones expressing adequate amounts of luciferase reporter in a tet regulated manner. Luciferase levels in two tagged, well regulated mES cell clones and a clone carrying the autoregulated construct integrated in the ROSA26 reference locus revealed strong variations that seem to be specific to autoregulation. *In vivo* bioluminescence analysis of transgenic mice derived from the two above mentioned clones showed a strong heterogeneity of expression between individuals of the same line with some animals completely shutting off expression. As expected for the autoregulated design, luciferase expression in one transgenic mouse line could only be induced by doxycycline if the animal had previously shown detectable basal levels of expression. This induction was fast (1-4 days) and reversible. For both transgenic lines activity of the cassette was restricted to certain tissues which was probably due to the nature of the respective integration site. This may also be the cause for lack of regulation in the second mouse line.

In conclusion, proof of principle for the envisioned "tag-and-target" strategy has been provided by this work and the tools and methods have been established to start large scale screening. Autoregulated, tet dependent expression from predefined, RMCE reusable loci in mES cells will allow the rapid and simple generation of mice with conditional transgene expression.

1 Introduction

1.1 Regulated transgene activity in mammalian cells and transgenic mice

Expression of transgenes in mice is a fundamental technology applied to answer biological questions as well as an essential model of human diseases and for gene therapy approaches. In a multitude of situations tightly controlled conditional activity of the introduced transgene is required. For instance, if a transgene's constitutive expression disturbs embryogenesis it will be impossible to elucidate its functions in adult mice. Consequently, a number of strategies have been developed over the last 20 years that allow to control expression levels or activity of transgene products in cell culture as well as in animals.

A general principle commonly exploited to obtain an inducible transgene in mice is genetic regulation (mostly by application of the Cre/loxP system). Here, transcription is blocked by a stop sequence flanked by recombinase recognition target sites. The transgene will hence only be transcribed after the recombinase has catalysed excision of this stop cassette (e.g. applied by Belteki *et al*., 2005; Hitz *et al*., 2007). A major drawback of this strategy, however, is that once transcription has been activated it cannot be downregulated again.

Reversibility is accomplished by drug inducible/repressible systems that operate either by modulating activity of the protein (post-transcriptional regulation) or by controlling transcription of the transgene (transcriptional regulation). The ideal regulation system should fulfil the following demands (Fussenegger 2001):

A) Specificity: In early attempts to regulate transgene expression endogenous regulatory elements were utilised that were activated by exogenous or stress signals (e.g. heat, hypoxia, metal ions). However, this strategy interfered with (stress) response mechanisms from the host cell leading to unpredictable effects. Thus, the regulation system should respond to heterologous or modified endogenous inducing/repressing molecules that are not toxic and do not crossreact with host regulatory networks avoiding pleiotropic side effects.

B) Inducibility: Low basal activity and high expression levels upon induction are favoured. The extent of inducibility is measured by the regulation factor (expression level after induction/basal expression level).

C) Bioavailability of the drug: For *in vivo* applications the inducer/repressor should rapidly penetrate all cells and tissues including blood-brain and placental barriers.

D) Reversibility: The system should allow repeated induction and repression of transgene expression by addition or depletion of the regulating drug. The time that is required to reach basal/maximal expression levels after aministration/removal of the respective drug is dependent on the degradation rate of the chemical (*in vivo* high pharmacokinetic turnover of the drug in all tissues is desired).

E) Immunogenicity: Components of the regulatory system should not evoke any immunogenic response from the host when applied *in vivo*.

F) Dose-dependence: Expression levels of the regulated transgene should proportionally correlate to concentrations of the regulating agent.

Post-transcriptional regulation

Control over a protein's activity may be gained by fusion of the transgene to a steroid-binding domain (SBD). The most common domains used for this purpose are derived from the glucocorticoid and oestrogen receptors (Picard *et al*., 1988; Mader *et al*., 1993). In the absence of the respective hormone, the large heat shock protein 90 (Hsp90) complex binds to the SBD leading to inactivation of the fused protein (e.g. by cytosolic retention in the case of a transcription factor). Addition of the steroid induces an allosteric change by binding to the SBD. Hsp90 is then released rendering the SBD fusion active. Major disadvantages of these systems include pleiotropic side effects of the inducing hormone and - *vice versa* - leakiness of repression due to endogenous hormone signalling of the host. However, this could partly be overcome by the development of a mutated oestrogen binding domain that does not recognise 17β-oestradiol but solely responds to 4-hydroxytamoxifen (Feil *et al*., 1997). Another factor impairing general application of this approach is that activity of SBD fusions proteins may differ from that of the original proteins.

Transcriptional regulation

Drug inducible transcriptional regulation systems consist of two components: 1) a transactivator/transrepressor initiating or blocking transcription dependent on presence or absence of a certain molecule (e.g. an antibiotic) and 2) a responsive promoter driving expression of the gene of interest (GOI) (reviewed by May *et al*., 2006).

The first transcriptional regulation system applied in mammalian cells was based on components of the *E. coli lac* operon (Hu and Davidson, 1987, Figge *et al*., 1998). The Lac repressor (LacI) expressed in the cells binds to its cognate operator sequence (*lacO*) in the DNA. If *lacO* is located in a promoter region, LacI will silence expression of the downstream

gene. Repression is abolished by the allolactose analogue IPTG whose binding diminishes affinity of LacI to *lacO*.

Another approach to regulate gene expression employs chemically induced dimerisation (CID). DNA binding and transactivating domains of the transactivator are separately expressed as fusions with a dimerisation domain that binds a certain drug. Upon addition of this drug the fusion proteins form heterodimers thereby constituting the functional transactivator. This principle is carried out in the FK506 binding protein 12 (FKBP)/FK506 or rapamycin and cyclophilin/cyclosporin regulation systems (Spencer *et al*., 1993).

Prominent regulated expression systems currently in use are controlled by dimerisation of transactivator homodimers or their conformational change induced by certain small molecules (mostly antibiotics). This determines the transactivator's capacity to bind to its target DNA sequence. Examples for this kind of regulation mechanism are the streptogramin system (Pip/pristinamycin; Fussenegger *et al*., 2000), the macrolide based system (E/erythromycin, clarithromycin, and roxithromycin; Weber *et al*., 2002) and most notably the tetracycline (tet) regulated system which will be described in detail later on (Gossen and Bujard, 1992). Combined use of these systems to create artificial regulatory networks allowed for sophisticated fine tuning of expression (Kramer *et al*., 2003). Also, independent expression of multiple genes can be managed by non interacting regulation systems, which is of interest e.g. to elucidate effects of different gene products in cascades. To extend the scope of drug regulated gene expression systems available, the cumate, the coumermycin/novobiocin and the TraR based systems have been developed recently (Mullick *et al*., 2006, Zhao *et al*., 2003, and Neddermann *et al*., 2003, respectively). These take advantage of regulatory elements from bacteria and viruses alike the previously described systems, a strategy that avoids possible interactions with host regulation networks.

Only a limited set of the systems described above is available for reversible transcriptional regulation in transgenic mice. Neither streptogramin nor macrolide regulated expression in transgenic mice has been established so far and the same is true for the novel cumate, coumermycin/novobiocin and TraR systems. CID employing FKBP or cyclophilin suffers from the immunosuppressive activity of the inducers FK506 or rapamycin and cyclosporin, respectively. Initial attempts to exploit the *lac* system in mice failed due to silencing of lac genes or the tendency of bacterial sequences to be excised from the host (reviewed by Scrable, 2002). No more than a few years ago, efforts by Cronin *et al*. proved that the *lac* system is functional in transgenic mouse lines (Cronin *et al*., 2001; Cronin *et al*., 2003, Ryan and Scrable, 2004). Although *lac* is a promising tool to regulate transgenes *in vivo*, the tet

system has a considerable head start as earliest reports of its application in mice reach back to the mid nineties (e.g.: Furth *et al*., 1994; Hennighausen *et al*., 1995; Kistner *et al*., 1996). Thus, the well established tet system still is the method of choice to govern reversible transgene expression in mice.

1.1.1 Tetracycline (tet) regulated gene expression

The tet system for regulated gene expression in eukaryotic cells was developed from the *E. coli* tetracycline resistance operon. There, the tet repressor (TetR, Postle *et al*., 1984) homodimer senses the presence of the protein synthesis inhibiting antibiotic. Upon binding of tet, TetR changes its conformation and does not recognise its operator sequence (tetO) in the DNA anymore (Müller *et al*., 1995). Thereby repression of the *TetA* gene is abolished, which codes for a membrane transporter that removes tet from the cell (overview e.g. by Hillen and Berens, 1994).

After studies had shown that this principle could be exploited for regulated gene expression in plants (Frohberg *et al*., 1991; Gatz and Quail, 1988, Gatz *et al*., 1991), Gossen and Bujard adapted it for the use in mammalian cells (Gossen and Bujard, 1992). The regulation mechanism of the so called Tet_{OFF} system is depicted in figure 1.

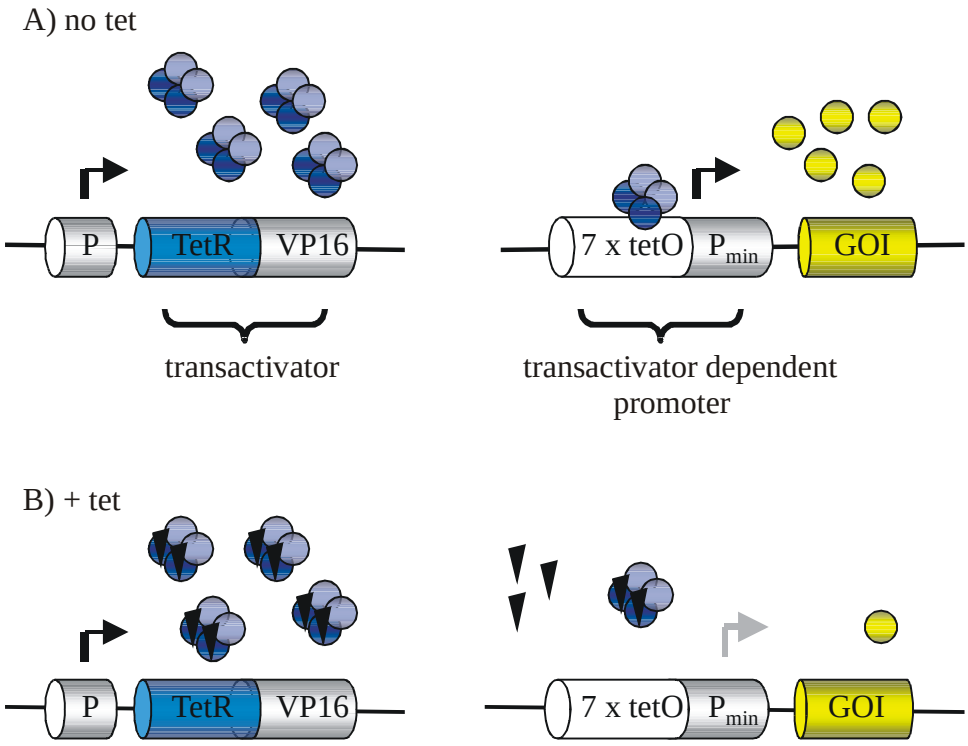


Figure 1: The Tet_{OFF} system mode of operation
A) The transactivator tTA consisting of the TetR repressor domain fused to the VP16 transactivation domain is constitutively expressed and homodimerises. In the absence of tet or its analogues (doxycycline or anhydrotetracycline), the repressor domain binds to the tet operator sequences (tetO) proximal of the minimal promoter $P_{hCMV*-1}$. The transactivating domain triggers expression of the downstream transgene. B) When tet (black triangles) is added it is bound by tTA which in turn changes its conformation. Thereby, binding of tTA to tetO is blocked and expression is shut off.

In the Tet_{OFF} system the transactivator tTA is constitutively expressed. It harbours a TetR repressor domain recognising tetO and a Herpes simplex virus VP16 transactivation domain (Triezenberg *et al.*,1988 a and b; Sadowski *et al.*, 1988). Homodimers of this fusion protein bind to the $P_{hCMV*-1}$ promoter which consists of 7 tetO repeats close to the human Cytomegalovirus minimal promoter P_{hCMV} (Boshart *et al.*, 1985). Thereby, the VP16 domain is capable of initiating transcription of the downstream transgene by recruitment of endogenous transcription factors like TFIIB (Lin *et al.*, 1991), TBP (Ingles *et al.*, 1991), or $TAFII_{40}$ (Goodrich *et al.*, 1993). Addition of tet or its analogues, e.g. doxycycline (dox) or anhydrotetracycline leads to a conformational change of tTA rendering the protein unable to bind to the promoter. Consequently, the transgene ceases to be transcribed.

1.1.1.1 Modifications and improvements of the tet dependent transactivator

Since its first application in mammalian cells the transactivator of the tet system has been modified to improve performance and to broaden its abilities.

Gossen *et al*. (1995) found that when only 4 amino acids in the TetR domain of tTA were changed the phenotype of the molecule was inverted (Tet_{ON}). This reverse transactivator (rtTA) binds only to tetO in the presence of tet. However it has to be noted that rtTA is only marginally activated by tet but well inducible by its analogue dox. Application of tTA is advantageous when a transgene should be active and needs to be quickly repressed while rtTA allows a rapid induction of expression.

One pitfall of the early tTA version lies in the viral VP16 transactivation domain which can induce large-scale chromatin decondensation and interacts with a multitude of transcription factors (Tumbar *et al*., 1999; transcription factor interactions reviewed by Flint and Shenk, 1997). High amounts of this molecule lead to an overall depletion of transcription factors needed for host RNA production which can result in cell death. This process is called "squelching" and was first observed when overexpressing Gal4 in yeast (Gill and Ptashne, 1988). To reduce adverse effects that were observed with high concentrations of (r)tTA (Gossen and Bujard 1992; Damke *et al*., 1995; Saez *et al*., 1997; Gallia and Khalili, 1998; Strathdee *et al*., 1999), the VP16 domain was substituted for different fusions of modified 12 amino acid minimal transactivation domains derived of VP16 (Baron *et al*., 1997). A motif consisting of 3 so called F domains strongly reduced negative effects (reduction to 33% of full VP16 domain) while retaining almost the full transactivation activity (98%); this tTA version was called tTA2. Higher tolerance to tTA2 is probably due to lack of multiple binding sites for transcription factors in comparison to the initial VP16 sequence. For instance, Oct-1 (Hayes and O'Hare, 1993), $TAFII_{40}$ (Goodrich *et al*., 1993), and ADA2 (Silverman *et al*., 1994; Candau *et al*., 1996) are not capable of binding to the minimal transactivation domain repeats. Alongside to tTA also rtTA has been modified by substituting the full VP16 domain for repeats of minimal domains (Kämper *et al*., 2002). In addition, other modifications like removal of potential splice sites and sequences that could form hairpin structures from synthetic VP16 domains improved performance of these rtTA versions (Urlinger *et al*., 2000a).

An alternative approach to avoid potential negative effects of VP16 was to use non viral transactivation domains, e.g. p65 (Schmitz and Bäuerle, 1991) and E2F4 (Ginsberg *et al*., 1994). Fusions of p65 with TetR proved to be as efficient as the VP16 transactivator while

E2F4 has a lower transcriptional activation potential (Urlinger *et al*., 2000b; Akagi *et al*., 2001).

Transcription can not only be induced by the tet system but also be efficiently silenced when TetR is fused to the KRAB (Krüppel-associated box) repressor domain of e.g. the human Kox1 protein (Margolin *et al*., 1994; Deuschle *et al*., 1995). The resulting molecule is accordingly called tTS (transcriptional silencer) while the reversely active transrepressor is termed rtTS (Hayakawa *et al*., 2006).

Since mutually exclusive TetR dimerisation domains derived from different gram-negative bacteria strains have been developed (Rossi *et al*., 1998; Baron *et al*., 1999), tTA, rtTA, and tTS may be used in the same cell avoiding formation of non active heterodimers. Alternatively, expression of two molecules of the transrepressor or of the transactivator linked by a small amino acid linker enforces homodimerisation (single chain transactivators/transrepressors; Krueger *et al*., 2003; Hayakawa *et al*., 2006). Combination of compatible tet regulating elements is for example exploited to reduce basal levels by using rtTA in combination with tTS: in the absence of tet, tTS binds to the tet dependent promoter and represses minimal expression. When tet is added tTS is displaced by rtTA which induces transcription (view e.g. Freundlieb *et al*., 1999; Förster *et al*., 1999; Bornkamm *et al*., 2005; Krueger *et al*., 2006).

Performance of transactivators is not only determined by the transactivation domain but also by their ability to translocate to the nucleus. Thus, fusion of a nuclear localisation sequence (nls) to tTA (Yoshida and Hamada, 1997) or rtTA (Gossen *et al*., 1995) improved the induction/repression potential of the transactivators by increasing the nuclear transfer rate.

Also, binding affinity of the transactivators to the tetO sequences in the responsive promoter is a determinant factor (Hinrichs *et al*., 1994; Orth *et al*., 1998; Baron *et al*., 1999). By subjecting the TetR domain to random mutagenesis rtTA variants could be identified that displayed enhanced DNA binding potential (Urlinger *et al*., 2000a).

1.1.1.2 Application of tet dependent transgene expression (in transgenic mice)

Tet regulated transgene expression has been utilised in a vast number of experimental approaches in different mammalian cell lines including HeLa, HEK293T and CHO. Regulation factors and basal expression levels are dependent on the individual cell line, indicating that cell specific factors influence performance of the system (Howe *et al*., 1995; Leuchtenberger *et al*., 2001).

Also murine embryonic stem (mES) cells can be manipulated by tet regulated transgene expression, facilitating studies concerning the intricate balance between self-renewal and differentiation of these pluripotent cells. For example, Masui *et al*. (2007) investigated the mechanism of Sox2 maintained pluripotency by tTA governed Sox2 expression in otherwise Sox2 null mES cells. In contrast, Kyba *et al*. (2002) used rtTA regulated HoxB4 expression to enforce hematopoietic differentiation of mES cells.

With the help of mES cells harbouring tet regulated transgenes or by pronuclei injection of tet regulated constructs into oocytes it is possible to generate mice that express/repress certain transgenes upon administration of dox (e.g.: Furth *et al*., 1994; Hennighausen *et al*., 1995; Kistner *et al*., 1996). Today, an increasing number of mouse strains expressing tTA or rtTA under control of different promoters allows to generate transgenic mice with tissue specific transgene expression by mating with mice that carry the GOI driven by a tet dependent promoter. This strategy permits spatial (tissue specific tTA/rtTA) as well as temporal (administration of dox) control of transgene expression. A table listing available tTA, rtTA and responder mouse strains is provided at: http://www.zmg.uni-mainz.de/tetmouse/.

Application of the tet system in transgenic mice for example enabled investigation of the circadian clock. Hong *et al.* (2007) generated mice carrying the *Clock* gene or a mutant dominant negative allele thereof under control of brain specific tTA. They observed discrete changes of circadian locomotor rhythms that reversed to the wild type phenotype upon administration of dox. Ernst *et al.* (2007) set out to devise a small animal model for Hepatitis C virus (HCV) infection by putting the HCV core protein or the complete ORF under control of liver restricted rtTA. Expression of the viral gene(s) can thus be switched on after maturation of the immune system circumventing immune tolerance to the viral protein(s). However, the authors detected significant transgene expression in the kidneys of induced mice. In addition, mosaic expression led to variable expression levels in individual mice of the same line implicating that the model still needs to be improved.

The tet system has also been used in combination with the Cre recombinase, for instance by controlling Cre expression by activity of tTA/rtTA (Schönig *et al*., 2002) or by subordinating rtTA expression to Cre mediated stop cassette excision (Belteki *et al*., 2005, Yu *et al*., 2005).

A new strategy does not employ the tet system for regulated expression of transgenes but for conditional knockdown of endogenous genes by RNA interference (RNAi, e.g. Matthess *et al*., 2005; Szulc *et al*., 2006). Seibler *et al.* (2007) placed tetO downstream of the U6 or H1 polymerase III promoters commonly used for shRNA expression. TetR was constitutively expressed and hence blocked transcription. Upon dox administration TetR detached from the

DNA allowing the shRNA coding sequence to be transcribed. By using an insulin receptor specific shRNA it was possible to induce severe hyperglycemia in transgenic mice within 7 days of dox administration. When a microRNA based shRNA is used, RNAi can also be controlled by conventional tet dependent polymerase II promoters. This allowes mating of mice carrying RNAi responder constructs to already existing (r)tTA strains. Dickins *et al*. (2007) employed this strategy to achieve tissue specific, reversible knockdown of the tumor suppressor gene p53.

A major concern of regulated expression in particular in transgenic mice is the basal activity in the repressed state, i.e. the system's "leakiness". Indeed, the tet system displays varying degrees of basal expression *in vivo* due to residual activity of the minimal, tet dependent promoter (reviewed by Sun *et al*., 2007). This may be tolerated if basal levels of the respective transgene are too low to elicit a phenotype. In case tightness of the tet system is critical, basal expression may be suppressed by the combined use of rtTA and tTS as described by Zhu *et al*. (2001). The authors had observed inflammation and other evidence for leaky expression in the lungs of mice harbouring an interleukin 13 (IL-13) transgene under the control of lung specific rtTA. However, this could be circumvented by simultaneous lung specific expression of tTS. While this did not influence inducibility of the rtTA regulated transgene, the phenotype in the repressed state was efficiently eliminated.

In conclusion, the tet system has been successfully applied in a number of transgenic mouse models to date. Its most important advantage is provided by the reversibility of transgene expression, overcoming a major limitation of classical (genetic) regulation systems.

1.1.1.3 Autoregulated, tet dependent expression

The basic tet system consits of two separate units: a constitutively active promoter driving (r)tTA and a responsive promoter controlling expression of the transgene. As constant high levels of transactivator have been shown to have cytotoxic or growth arresting effects probably due to squelching (e.g. Gossen and Bujard, 1992; Damke *et al*., 1995; Saez *et al*., 1997), autoregulated expression systems have been developed where levels of transactivator in the uninduced state are significantly reduced. In these systems, (r)tTA governs its own expression. Different strategies to achieve autoregulated expression are outlined in figure 2. Initiation of expression in these systems is dependent on a certain basal expression providing enough transactivator molecules to trigger transcription. After activation of transcription, a positive feedback loop of expression is generated.

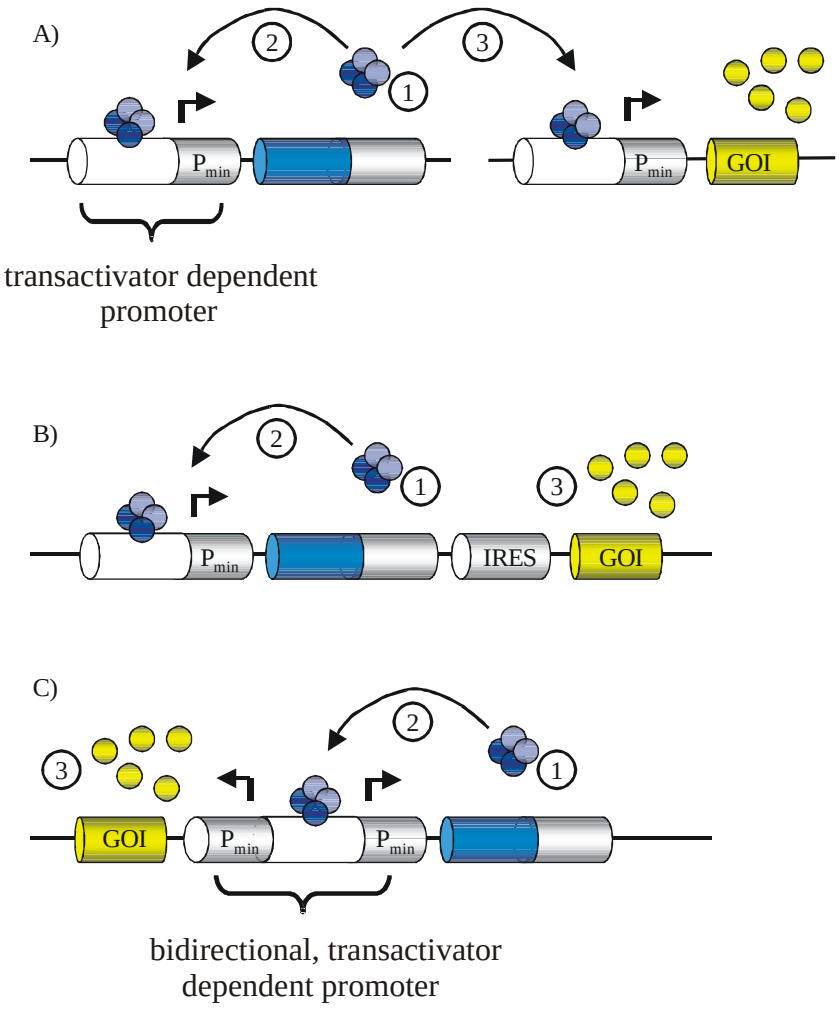


Figure 2: Different modes of autoregulated expression
1) The transactivator dependent promoter drives expression of the transactivator. 2) The transactivator binds to the operator sequence of the promoter and in turn triggers expression of more transactivator. 3) At the same time, expression of a transgene is induced. A) Dual promoter system: transactivator and transgene are driven by their own, transactivator dependent promoter. These two expression units may reside on separate or on the same plasmid. B) The transactivator dependent promoter drives expression of a bicistronic mRNA. Translation of the second cistron is mediated by an internal ribosomal entry site (IRES). C) A bidirectional transactivator dependent promoter governs transcription of both transactivator and transgene.

Initial attempts to apply autoregulated tet systems showed cellular toxicity (Gallia and Khalili, 1998; Strathdee *et al.*, 1999). However, this was most likely not specific for autoregulated expression but due to the use of old transactivator variants harbouring the full VP16 transactivation domain. Autoregulated tet systems and their expression charasteristics have been of interest in a number of studies (e.g. Hofmann *et al.*, 1996; Unsinger *et al.*, 2001; Unsinger *et al.*, 2004). A prominent advantage of autoregualted cassettes designed as depicted in figure 2 panels B and C is their compactness that facilitates single step transduction, e.g. by viruses.

For instance, viral transduction of tet autoregulated therapeutic genes has been envisaged for gene therapy approaches (Gould *et al*., 2000; Unsinger *et al*., 2001; Chtarto *et al*., 2003). Kühnel *et al*. (2004) infected mice with tTA4 autoregulated retroviruses demonstrating that this system facilitates inducible gene expression in the liver *in vivo*. By implementing autoregulatory rtTA driven transgene expression in lentiviruses, Markusic *et al*. (2005) could show that the autoregulatory loop was superior to constitutive rtTA expression: 1) viral titers were higher, 2) rtTA and the transgene were not detectable in the absence of dox, 3) induction kinetics were improved, and 4) induction levels were increased.

The only reports applying an autoregulated tet system in transgenic mice so far have been published by Shockett *et al*. (1995; 2004). There, the authors controlled the recombination activating genes 1 and 2 (RAG-1 and RAG-2) by an autoregulated tTA as depicted in figure 2 panel A (dual promoters). In mice with a RAG-1 or RAG-2 KO background induction of both transgenes partially rescued V(D)J joining, reconstituting lymphocyte function to a certain degree.

1.1.1.4 Expression characteristics of the autoregulated tet system

When the transcativator is abundant in cells due to constitutive expression, tet systems gradually increase transgene expression with increasing (rtTA)/decreasing (tTA) tet concentrations (Baron *et al*., 1997; Kringstein *et al*., 1998; Bornkamm *et al*., 2005). Autoregulated systems in contrast display a bimodal response: cells either display basal transgene expression or they express maximal levels once a certain level of inducer is reached (Becskei *et al*., 2001; May *et al*., 2007, submitted). With increasing concentrations of dox (referring to rtTA), cells do not express higher amounts of the transgene but rather the probability of expression in individual cells is augmented. This expression characteristic is typical for positive feedback loops (Becskei *et al*., 2001). It can be explained by different molecular conditions in binary and autoregulated Tet_{ON} systems: When the transactivator is constantly expressed, the degree of regulated expression is only limited by the the amount of inducer present. In the autoregulated setting, inducer as well as rtTA are limiting factors for initiating transcription.

1.2 Methods for introducing transgenes in mice

A prerequisite for the generation of transgenic mice is the integration of the GOI and - if desired - regulatory elements into the chromosomes of either mES cells or fertilised oocytes. Expression of these cassettes strongly depends on the nature of the integration site, a phenomenon which is called the "position effect". Possible influences by endogenous regulatory elements neighbouring the integration site(s) or by heterochromatin formation include tissue or cell type restricted expression of the GOI, elevated basal expression levels or - very often - complete silencing of the GOI. An extreme example for hetreogenous expression due to position effects is provided by Feng *et al*. (2000): here, 24 different founder animals displayed 24 different expression patterns.

Several methods are available today that overcome the difficulties inherent to random integration. These include targeted integration into defined loci by either homologous or site specific recombination or shielding the transgene by large DNA sequences employing e.g. Bacterial Artificial Chromosomes (BACs). In the following chapters, these methods will be described and compared concerning their advantages and disadvantages for the generation of transgenic mice.

1.2.1 Homologous Recombination

Homologous recombination (HR) is one of the key mechanisms to create genetic diversity in nature. It mediates sequence exchange between chromosomes, thus allowing new combinations of alleles to be tested for their performance and to separate advantageous from unfavourable alleles.

HR is employed as a tool to specifically manipulate loci in living cells – in particular mES cells. By transduction of a user-defined DNA sequence flanked by homology arms into the cells, this sequence can be introduced into the homologous locus. Thus, gene function can be completely abolished (knock-out), modified (e.g. by creating hypomorphic alleles) or new transgenes may be introduced (knock-in). Although HR was first utilised in differentiated cells (Smithies *et al*., 1985), its frequency in this background is very low (10^{-9} – 10^{-8} events/cell) (Puttini *et al*., 2005) and it is often masked by the dominance of illegitimate recombination. However, attempts are made to alleviate application of site-specific HR in highly differentiated cells by stimulation with the meganuclease I-SceI (Puttini *et al*., 2005).

Murine ES cells in contrast display a much higher probability of HR (10^{-6} – 10^{-5} events/cell) and once a gene has been correctly targeted, the probability that a second fragment is integrated randomly is approximately 20 fold decreased (Reid *et al*., 1991).

Efficiency of HR in mES cells is dependent on the length of homology arms in the targeting construct: maximal efficiency is reached at a total length of 14 kbp while homologous arms shorter than 1 kbp each are not sufficient for HR (Thomas *et al*., 1992). Also, targeting efficiency by HR is influenced by the individual locus. Screening for correctly targeted cell clones can be facilitated by adapted selection protocols. ROSA26 (Friedrich and Soriano, 1991; Zambrowicz *et al*., 1997), a locus that is commonly used for knock-in strategies for instance can be targeted with an frequency of up to 25% (personal communication J. Haigh).

As mES cells can be used to generate transgenic mice, e.g. to elucidate gene function during ontogenesis or for modeling human diseases, use of HR in mES cells has revolutionised basic and biomedical research (reviewed by Downing and Battey, 2004).

1.2.2 Site specific recombination

Next to HR, also site specific recombination allows to manipulate mammalian genomes. Site specific recombination systems do not require any endogenous factors and consist of two basic components: the site specific recombinase (SSR) enzyme and its cognate recognition target (RT) site in the DNA. The most common SSRs currently used are the Cre and Flp tyrosine recombinases, which share a common recombination mechanism: DNA is cleaved, exchanged and ligated at the specific RTs (Sadowski *et al*., 1995). Cre was derived form the bacteriophage P1 (Sternberg *et al*., 1986) and recognises so called loxP (locus of crossover) sites whereas Flp was found in the yeast 2µ circle and recombines FRT (Flp recognition target) sites. While both types of RTs differ in their nucleotide sequence, their overall structure is similar: an asymmetric spacer sequence is flanked by two 13 bp inverted repeats (figure 3).

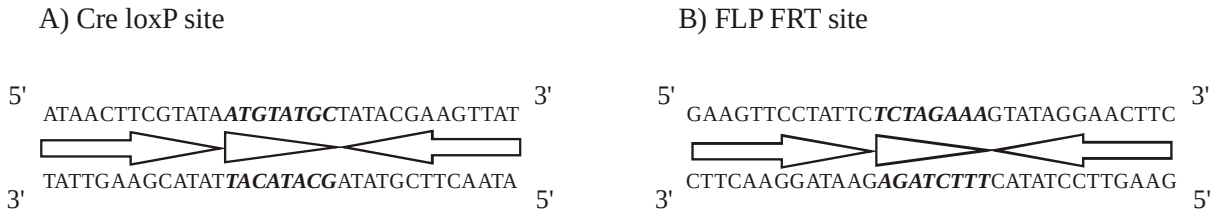


Figure 3: Sequences of lox P and FRT sites
RTs of both Cre and Flp share a common structure: an asymmetric 8bp core or spacer sequence (indicated by triangle) is flanked by inverted 13 bp repeats (marked by arrows). A) Cre recombinase RT loxP B) Flp recombinase RT FRT

Recombinase monomers bind to their respective RTs promoting formation of a DNA synaptic complex and recombination between two identical RTs in the same orientation (Hoess *et al*., 1985a and b; Amin *et al*., 1991). Thus, spacer asymmetry determines the outcome after recombination of varying positions and orientations of RTs (figure 4). DNA flanked by inverted RTs will be inverted by the SSRs (figure 4 panel A) whereas direct repeats lead to excision of a circular molecule of flanked DNA (figure 4 panel B). The opposite reaction, namely integration of circular DNA carrying an RT into the linear molecule, can also occur albeit with a much lower frequency due to thermodynamic reasons (figure 4 panel B). RTs positioned on separate DNA molecules result in translocation of distal sequences (figure 4 panel C).

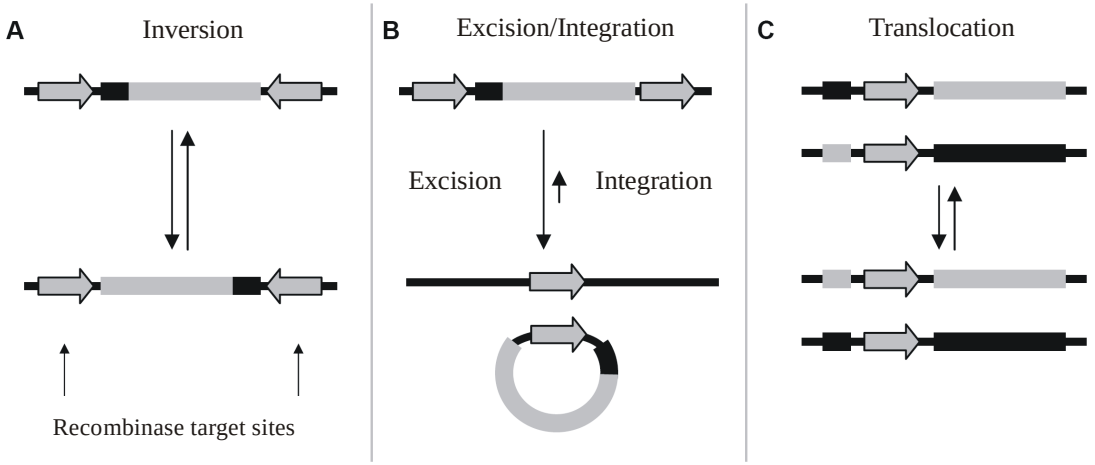


Figure 4: Outcome of recombination reactions mediated by Cre and Flp
RTs of Cre and Flp and their respective orientation are indicated by arrows. A) Inverted RTs on the same molecule lead to inversion of the flanked DNA sequence after recombination. B) Direct repeats of RTs on the same molecule result in excision of a DNA circle. Excision is much more efficient than the integration reaction as indicated by arrows. C) RTs on different DNA molecules result in DNA exchange of part of the linear molecules.

Cre recombinase works satisfyingly in mammalian cells in spite of its prokaryotic codon bias; codon optimisation yielded an enzyme version termed iCRE, which showed only slightly improved performance (Shimshek *et al*., 2002). Wild type Flp in contrast suffered from low activity at 37°C (Buchholz *et al*., 1996), a drawback that was overcome by generating a mutant with an adequate optimum temperature termed Flpe (enhanced Flp; Buchholz *et al*., 1998). Recently, a new mutant of Flpe recombinase called Flpo (optimised Flp) has been published, which is claimed to be more efficient for recombination due to codon optimisation for mammalian cells (Raymond and Soriano, 2007).

Both SSRs were further refined by fusion to an oestrogen receptor ligand binding domain (LBD), rendering them inactive in the absence of ligand and hence attaining temporal control over SSR activity. The LBD version employed is the ER(T2) mutant (chapter 1.1), which solely binds 4-hydroxytamoxifen (CreER(T2): Feil *et al*., 1997; FlpeER(T2): Hunter *et al*., 2005).

Cre and Flpe are in particular valuable tools for genetic analyses in mice (reviewed by Branda and Dymecki, 2004; Schnütgen *et al*., 2006). They enable conditional gene expression by removal or inversion of a transcriptional stop cassette, conditional knockout of targeted genes, removal of undesired selection markers and Recombinase Mediated Cassette Exchange (RMCE), which is described below. Today, a multitude of mouse strains are available that express Cre in a tissue specific and/or inducible manner ("Cre-zoo"; listed at: http://www.mshri.on.ca/nagy/) and the number of analogous Flpe lines is increasing (Farley *et al*., 2000; Rodriguez *et al*., 2000; Awatramani *et al*., 2003). These strains facilitate spatial and temporal control over transgene expression *in vivo* by simple mating to responder lines of interest.

A new Cre-like recombinase termed "Dre" was identified that is a new candidate for engineering mammalian genomes (Sauer and McDermott, 2004). Since it recognises its own specific set of RTs it could be used in parallel to Cre and Flpe thus expanding possibilities of genetic manipulation.

Another recombinase that has already been successfully applied in mammalian cells was derived from the phage ΦC31 (Thorpe and Smith, 1998; Andreas *et al*., 2002; Smith and Thorpe, 2002; Groth *et al*., 2000). ΦC31 belongs to the serine family of recombinases and follows a mechanism that is distinct from Cre, Dre and Flp (depicted in figure 5 panel B). This enzyme recombines the non identical attB and attP sites, inserting a DNA sequence and creating attL and attR sites. AttL and attR cannot recombine so that the integration reaction mediated by ΦC31 is essentially irreversible. ΦC31 has been proven to work efficiently in mES cells (Belteki *et al*., 2003) and for gene therapy approaches (Olivares *et al*., 2002; Ortiz-Urda *et al*., 2002) indicating its potential for future application.

1.2.2.1 Recombinase Mediated Cassette Exchange (RMCE)

Transgene expression levels strongly depend on the individual integration site of the transgene since neighbouring elements like silencers or chromosomal structures have a strong impact on transcription. Random integration (RI) of a GOI thus suffers from unpredictable, potentially unstable expression characteristics. This problem can be circumvented by HR; however, only a limited number of well described loci suitable for transgene expression is available so far and reasonably efficient use of HR is restricted to mES cells.

Utilisation of SSRs for Recombinase Mediated Cassette Exchange (RMCE) presents an alternative tool for targeted integration of transgenes that is applicable in all cell types. In addition, RMCE allows for more efficient selection strategies to identify correct integration events than HR.

A prerequisite for RMCE employing Cre and Flpe was the development of heterologous, non interacting loxP and FRT sites, respectively (reviewed by Branda and Dymecki, 2004). These mutant sites fall into two classes: spacer variants and inverted-repeat variants. Spacer variants carry substituted nucleotides in the 8bp core sequence. As long as the spacer length is not changed, recombination is possible but restricted to homotypic pairs of RTs (Hoess *et al*., 1986; Senecoff *et al*., 1988). As ΦC31 recombines the non identical attB and attP sites, the naturally occuring heterotypic sites may be used for RMCE in this case.

RMCE involves two basic steps:

1) "tagging" of a genomic locus by intergrating heterotypic RTs either via HR or RI. In the case of RI, expression characteristics of the randomly tagged locus can be evaluated by a co-integrated reporter gene to determine its usefulness for further approaches.
2) "targeting" of the tagged site allows to exchange the tagging cassette for the desired transgene flanked by the same set of RTs (depicted in figure 5).

Applying this technique, pre-defined chromosomal loci may be reused or modified circumventing extensive screening procedures (reviewed by Bode *et al*., 2000; Baer and Bode, 2001).

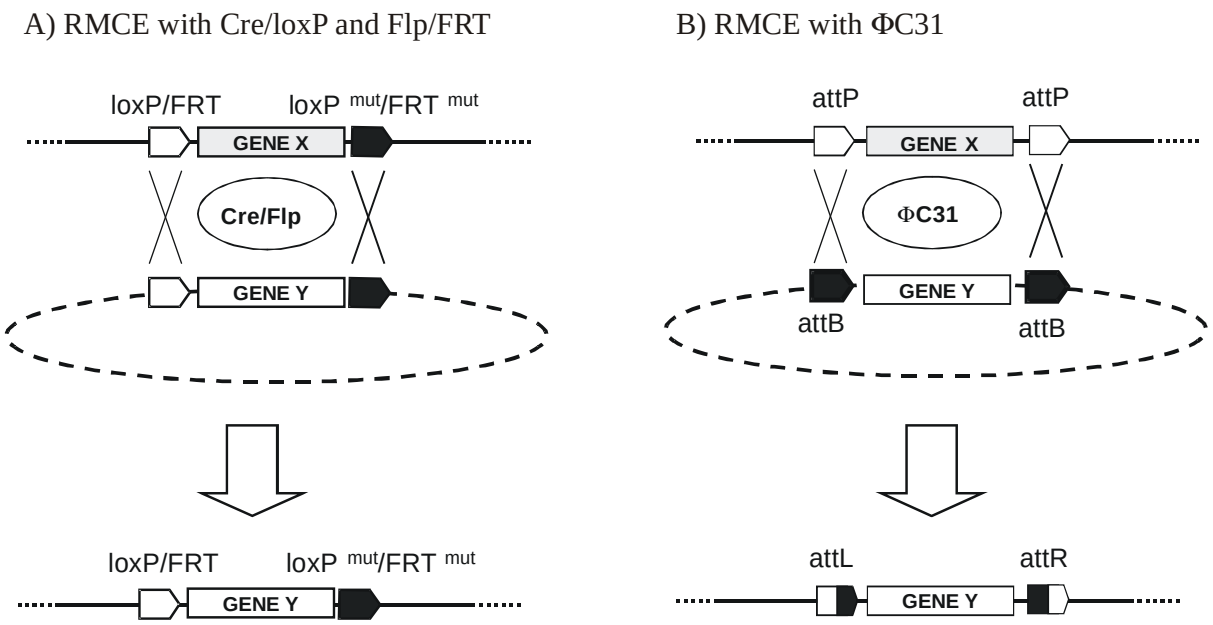


Figure 5 (from Wirth *et al.*, 2007): Principle of Recombinase Mediated Cassette Exchange (RMCE)
A) Flp and Cre recombinases recognise their specific target sequences (loxP and FRT, respectively) in the previously tagged genomic DNA and the incoming targeting vector. After integration of the targeting vector via either of the RTs and subsequent resolution of the intermediates formed, the genomic locus will have acquired gene y. Suitable selection strategies permit the recovery of the desired exchange event. Complete excision is prevented by using heterospecific (mutant) LoxP or FRT sites. Alternatively, inverted orientation of the RTs with respect to each other may be employed. P: promoter; the dashed lines indicate the prokaryotic vector backbone.
B) Site directed integration by ΦC31
The ΦC31 recombinase mediates recombination between the heterotypic attB and attP sites. A cassette flanked by two attB sequences is integrated into the genomic site and hybrid attL and attR sites are generated which are not compatible for any further recombination events. Thus, integration of the desired cassette is strictly unidirectional.

Several parameters influence efficiency of RMCE and by manipulating them the number of correct targeting events can be significantly enhanced.

For example, different site specific selection strategies have been devised to counterselect random integration of the targeting vector. One of these strategies is complementing a truncated version of a positive selection marker with its missing sequence from the incoming targeting plasmid and subsequent selection with the respective drug (e.g. complementing 5' Δ neomycin phosphotransferase with ATG, e.g. Schucht *et al.*, 2006; or reconstituting an HPRT minigene as in Wallace *at al.*, 2007). In a similar approach a promoterless positive selection marker is only expressed by a matching promoter from the parental locus after correct integration (Cobellis *et al.*, 2005). In addition, a negative selection marker such as the thymidine kinase gene may be included in the initial tagging construct, thus allowing to specifically eliminate all cells with the non exchanged cassette (selection with ganciclovir for absence of thymidine kinase; e.g. Toledo *et al.*, 2006; Wong *et al.*, 2005). A third possibility is to place diphteria toxin A in the non exchanged region of the targeting plasmid in order to

select against random integration into the genome (Araki *et al*., 2006). However, it has been reported that the ratio of correct RMCE events versus random integration may be acceptable even in the absence of any site specific selection (2 correct clones/30 analysed by Cobellis *et al*., 2005; 50% efficiency observed by Masui *et al*., 2005).

In order to perform RMCE, not only the targeting vector has to be present in the cell but also adequate amounts of the recombinase protein. Mostly, a second plasmid encoding for the recombinase is cotransduced (by transfection or electroporation) which leads to transient production of the protein. Alternatively, fusions of Cre (and Flp, unpublished data by Edenhofer *et al*.: Mouse Molecular Genetics Meeting, Heidelberg 2005) with the TAT sequence from HIV-1 have been created so that the enzyme can be directly administered to the cells and pass through the membranes of its own accord (Peitz *et al*., 2002). By retroviral pseudotransduction it is even feasible to obtain transient recombinase activity in a certain subset of cells dependent on the receptor specificity of the pseudovirus (Galla *et al*., 2004).

Transient activity of the recombinase is preferable in the case of Cre and ΦC31 because higher enzyme levels over an extended period of time have been shown to be harmful for the cells (Cre: Loonstra *et al*., 2001; Schmidt-Supprian and Rajewsky, 2007; ΦC31: Erhardt *et al*., 2006, Liu *et al*., 2006a). This is caused by the presence of cryptic loxP and attB/P sites in the genome that may be recognised and recombined leading to chromosomal rearrangements. To circumvent this problem - especially in Cre transgenic mice - , timely restricted Cre activity is utilised (described in chapter: 1.2.2). Of note, no overt adverse effects of Flp recombinase have been reported so far.

RMCE has been exploited in multiple fields ranging from basic reseach to biotechnological applications, e.g. for generation of recombinant virus producer cell lines providing high and predictable virus titers (Schucht *et al*., 2006; reviewed by Wirth *et al*., 2007). Endeavours to trap every gene in mES cells to generate (conditional) knockout mouse strains (Austin *et al*., 2004; Auwerx *et al*., 2004) have been made compatible with RMCE by including heterospecific RTs in the trapping vectors. Thus, trapped chromosomal loci in mES cells are readily accessible by RMCE (proof of principle e.g. by Cobellis *et al*., 2005). Available trapped mES cell clones are provided by the International Gene Trap Consortium (IGTC) at http://www.genetrap.org.

Recently, Wallace *et al*. (2007) impressively demonstrated the faculties of RMCE by replacing the murine α-globin regulatory domain for the syntenic human sequence. This exchange involved a segment larger than 100 kbp and produced an accurate mouse model for human α-thalassemia.

1.2.3 Bacterial Artificial Chromosomes (BACs)

Besides specifically targeting a transgene to predefined loci in order to avoid unwanted position effects, large DNA segments (up to 300 kb) in the form of BACs may be exploited to intercept influences from the host's genome. Since recombination systems were established that allow to manipulate BACs, transgenes may be modified if desired and introduced into the mouse germline via this route (reviewed by Giraldo and Montoliu, 2001; Sparwasser and Eberl, 2007). Either the GOI is kept in its endogenous surrounding thereby retaining potential regulatory elements that are not located in the immediate proximity, or it may be inserted into a BAC derived from any "neutral" locus (e.g. ROSA26: Giel-Moloney *at al*., 2007). The observation that expression levels are proportional to the integrated number of copies indicates that expression from BACs is largely liberated from position effects (e.g. Bender *et al*., 2007).

1.2.4 Comparison of HR, Site Specific Recombination, and BACs for generation of transgenic mice

The above described techniques to evade position effects of integrated transgene expression differ in their capabilities. HR employed to target endogenous loci for instance requires screening of a considerable number of clones since the gene targeting vector frequently integrates randomly. In addition, the large homology arms needed for efficient recombination hamper vector manipulation by simple cloning strategies. Another constraint of HR is the limited number of supportive loci available.

RMCE in contrast facilitates targeted integration with an efficiency of up to 100% (e.g. Schucht *et al*., 2006) due to apt selection strategies involving comparatively small constructs that are easy to handle. However, this method requires a previous integration step of the RTs by either HR or random integration.

BAC transgenesis on the one hand profits from the cointegration of sequences flanking the transgene thereby maintaining its cognate environment in the host. On the other hand, these sequences can as well contain other genes that may be coexpressed and potentially lead to unpredictable effects (reviewed by Matthaei, 2007).

1.3 In vitro differentiation of mES cells

mES cells were first successfully isolated from blastocysts and propagated *in vitro* in 1981 (Evans and Kaufman 1981; Martin, 1981). A prerequisite for mES cell culture was to identify conditions that support pluripotency, i.e. the potential to differentiate into cell types of all three germ layers. Maintenance of an undifferentiated yet pluripotent state of mES cells is dependent on factors present in the media, most notably leukaemia inhibitory factor (LIF) and BMP4 (Ying *et al.*, 2003). Removal of these factors induces spontaneous differentiation of the cells which may be directed to specific lineages by appropriate cues like cytokines.

Thus, manipulated mES may not only be exploited to establish mouse models but can also serve to model embryonic development *in vitro*. Besides its value for basic research, *in vitro* differentiation of ES cells may also be utilised as an inexpensive tool for drug testing (e.g. Reppel *et al*., 2007), in particular if teratogenicity is concerned. Since the isolation of human ES cells in 1998 (Thomson *et al*., 1998), *in vitro* differentiation of ES cells has attracted notice as a potential source for cells and tissues for the treatment of various diseases.

A multitude of *in vitro* differentiation protocols have been developed that allow generation of cell types of all three germ layers from mES cells, e.g. neuronal, cardiac, pancreatic and hepatic cells (reviewed by Wiese *et al*., 2006; Keller, 2005).

Pilat *et al*. (2005) differentiated mES cells *in vitro* to hematopoietic precursors that were almost indistinguishable from isolated bone marrow cells. These so called ES-HCs (ES cell derived hematopoietic stem cells) had a similar potential to reconstitute the hematopietic compartment in irradiated mice as adult hematopoietic stem cells (HSCs). This underlines that fate decision during *in vivo* differentiation may be reproduced with high accuracy *in vitro* if ES cells are subjected to adequate signalling.

1.4 Objective of this work

Random integration of transgenes into mammalian chromosomes mostly results in unpredictable and/or unstable expression patterns due to position effects. This in particular applies for regulated transgene expression. The aim of this work was to identify approaches that allow high and strictly controllable expression of Tet_{ON} regulated transgenes in mES cells for establishing mice with tightly regulated transgene expression. One approach should employ a single cassette harbouring an autoregulated reporter gene to screen for suitable loci in the mouse genome (“tagging”). In light of the fact that expression of genes governed in an autoregulated manner has not been investigated in detail on transgenic mouse level so far, a

central issue in this regard should be to characterise the performance of autoregulated cassettes and to evaluate their potential for further applications.

Transgene expression is frequently altered upon differentiation of mES cells to somatic cells, potentially resulting in tissue or cell type specificity or even complete silencing of the transgene in mice. Hence, another aspect should be to test strategies for evaluating expression after *in vitro* differentiation protocols to facilitate pre-screening for mES cell clones with suitable expression properties.

Potential reuse of tagged loci in such pre-characterised clones for further applications should be envisaged by implementing an approach applying site specific recombinases for cassette exchange.

On the whole, mES cell lines with reusable loci should be generated that provide high and well regulated transgene expression in transgenic animals. Time consuming screenings of mES cell clones as well as transgenic mice for adequate expression levels would be circumvented by the specific use of predefined, well characterised chromosomal loci with predictable expression characteristics.

2 Results

The objective of this work was to create mES cell lines that allow to establish transgenic mice with predictable, uniform and regulatable expression characteristics. For this purpose autoregulated expression modules based on the tet system were used, in which both expression of the gene of interest (GOI) and of the transactivator are controlled by doxycycline (dox). Since the performance of regulated gene expression is massively influenced by neighbouring sequences, the rationale was to characterise randomly tagged loci in mES cells by determining tet system regulated reporter activity. Subsequently, these loci may be reused by Recombinase Mediated Cassette Exchange (RMCE), a method that comprises two basic steps: 1) "Tagging" random loci with a reporter construct and 2) "targeting" sites conferring appropriate expression levels of the reporter gene with the GOI. Thereby it is possible to reuse a precharacterised locus.

2.1 Overall strategy for screening and reuse of chromosomal loci providing regulated expression

The overall strategy is depicted in figure 6. It includes the following steps: first, mES cells are transduced with an autoregulated reporter construct that also introduces FRT sites required for RMCE and a selection marker. After drug selection for stable integration, mES cell clones displaying low basal reporter gene expression and maximum expression levels when induced are singled out. Thus, established clones are not only screened for high expression levels but also for their ability to support adequate regulation. Appropriately performing clones are subsequently screened for the presence of only one copy of the tagging vector.

Gene expression from a given locus can change during the process of differentiation – it may be completely silenced in the transgenic animal or become restricted to certain tissues or cell types. Consequently, tagged integration sites have to be further evaluated by *in vitro* differentiation and teratoma formation as means to validate clones before generation of reporter mice. Finally, transgenic animals are established and analysed.

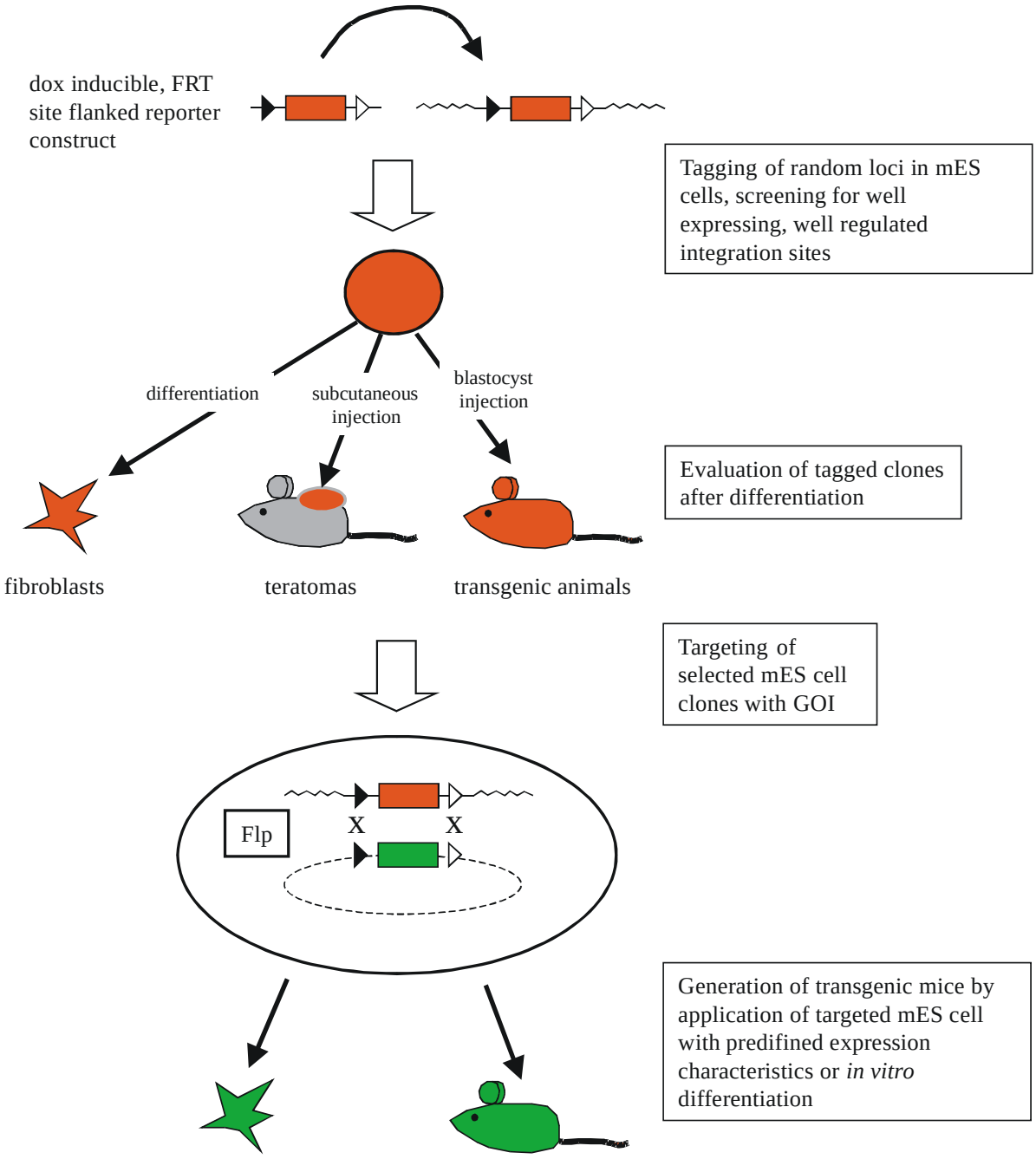


Figure 6: Overview strategy
Chromosomal loci in mES cells are randomly tagged with a dox inducible, FRT flanked screening cassette. Clones with high and well regulated expression are subjected to diverse differentiation approaches to evaluate their behaviour after commitment to mature cell types. ES cell clones characterised by these means are used for site specific integration of the GOI into the tagged locus via RMCE. Targeted mES cells can be directly used for the rapid generation of transgenic mice circumventing extensive screening procedures, for precharacterised loci admit predictable, high and well regulated transgene expression.

2.2 Tagging mES cells with an autoregulated, tet-dependent reporter cassette

To identify and retarget appropriate integration sites in mES cells, the pTagTK/NPT screening vector (figure 7) was employed which comprises the following elements: 1) firefly luciferase as a sensitive reporter controlled by the Tet_{ON} system in an autoregulated fashion, 2) heterologous FRT sites to reuse tagged loci via RMCE and 3) a fusion of thymidine kinase (*tk*) and neomycin phosphotransferase genes (*npt*) for selection of tagged clones.

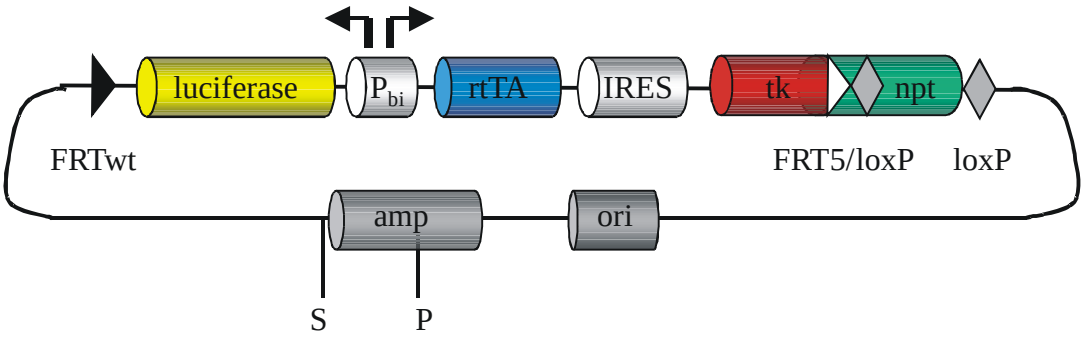


Figure 7: Schematic view of the tagging vector pTagTK/NPT
Two separate transcripts are produced from the bidirectional, tet dependent promoter (P_{bi}): One codes for firefly luciferase mutant *luc+* and the second comprises the reverse transactivator variant rtTA2(S)-M2 and a fusion of the thymidine kinase gene (tk) with the neomycin phosphotransferase gene (npt). Translation of this fusion is mediated by the Encephalomyocarditis virus internal ribosombal entry site (IRES). The FRT wild type and mutant FRT5 sites are indicated by black and white arrowheads, repectively.
grey diamonds: wild type loxP sites; amp = ampicillin resistance gene; ori = origin of replication; S = *SspI* restriction site; P = *PvuI* restriction site

Firefly luciferase mutant *luc+* (Promega) was chosen for the screening since it is a very sensitive reporter. Its expression is controlled by the bidirectional, tet dependent promoter (Baron *et al.*, 1995) which at the same time drives expression of the reverse transactivator (rtTA; variant rtTA2(S)-M2, Urlinger *et al.*, 2000a) of the tet system (Gossen and Bujard, 1992). This arrangement of the different components creates a regulatory positive feedback loop (figure 8): In the absence of the inducer doxycycline (dox), the tet dependent promoter is only marginally active allowing for a certain level of basal expression. When dox is added it is bound by the few molecules of rtTA present which will dimerise and bind to the bidirectional tet dependent promoter. This initiates transcription of the reporter and *rtTA*, hence promoting and enhancing transcription. The fact that transcription can only be started when the cassette is not completely silent without any dox has to be taken into account for the design of experiments with this system.

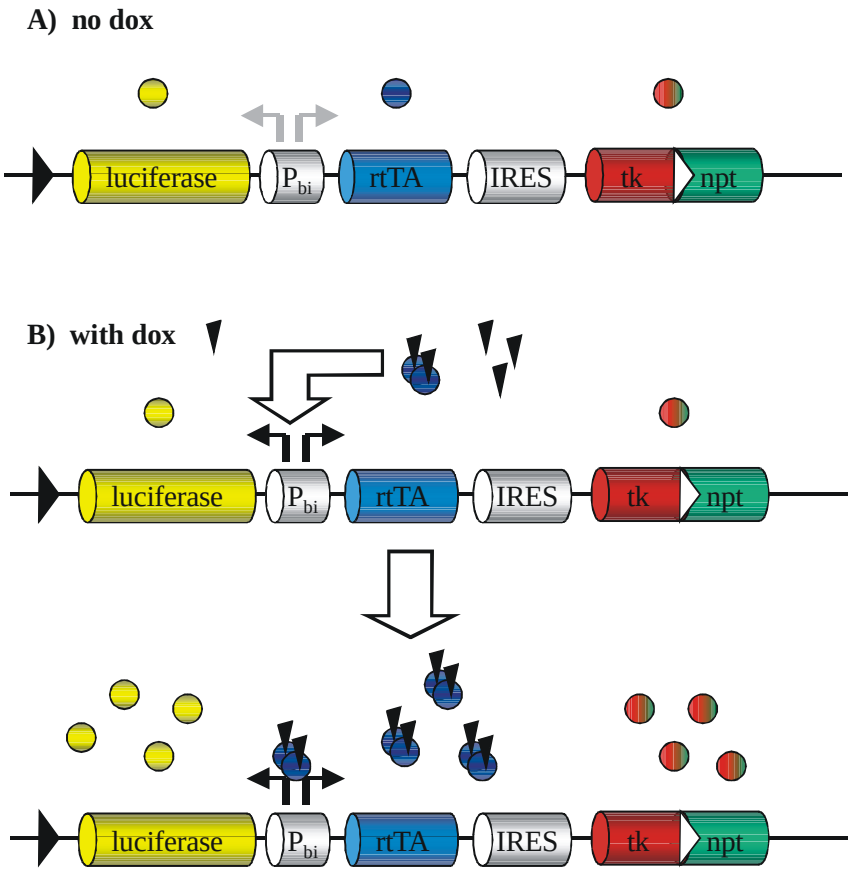


Figure 8: Autoregulated positive feedback loop
A) In the absence of dox basal activity of the promoter leads to the production a low concentration of rtTA.
B) When added, dox is bound by rtTA which in turn dimerizes, binds to the promoter and enhances transcription. A positive feedback loop is generated because the reverse transactivator triggers production of its own mRNA.

An Encephalomyocarditis virus (EMCV) internal ribsomal entry site (IRES) mediates translation of a fusion of the thymidine kinase gene (*tk*) with the neomycin phosphotransferase gene (*npt*). Thymidine kinase operates as a negative selection marker: it phosphorylates ganciclovir (GCV) which is toxic for the cells as a triphosphate. *Npt* is a positive selection marker for eukaryotic cells for it confers resistance to the selection drug G418. A wild type FRT site is placed next to the 3' end of the luciferase gene while the mutant FRT5 site is integrated in frame in-between the *tk* and the *npt* part of the fusion.

The non identical FRTwt and FRT5 sites flank the expression core unit and allow tagging of chromosomal loci and their subsequent reuse by RMCE (details for exchange strategy given in chapter 2.3). This specific pair of recombination target sites has been shown to work efficiently before (e.g. Schucht *et al.*, 2006).

In addition, loxP sites flanking *npt* allow to remove this selection marker (figure 7). As *npt* has been reported to have a negative influence on expression in some cases, this option was implemented. *Tk* expression leads to male sterility in mice (Cesari *et al.*, 2004; Braun *et al.*,

1990), but since animals are established without dox (i.e. in the off state) this should not impede generation of a mouse line.

2.2.1 Generation of mES cell clones tagged with pTagTK/NPT

For stable integration into mouse ES cells, pTagTK/NPT plasmid DNA was linearised with a restriction enzyme that cuts within or close to the ampicillin resistance gene. The digest left bacterial sequences flanking the actual tagging cassette (*PvuI*: ~850 bp/ *SspI*: ~470 bp next to the FRTwt site and *PvuI*: ~1,4kb/*SspI*: ~1,84 kb next to the polyadenylation signal of the *tk/npt* fusion). The bacterial backbone of the vector was not removed to protect the cassette from nucleases that would affect the integrity of the tag.

Low passage IB10 mES cells (subclone of E14 derived from 129/Ola mice; Robanus-Maandag *et al*., 1998) were electroporated under different conditions as given in table 1. The tagging step was done by electroporation, because this transduction method leads to a relative high percentage of cells with a single copy integration and does not interfere with germ line competence of the cells. Electroporated cells were induced with dox after one day and put under pressure with 0,4 mg/ml G418 two days after electroporation.

Table 1: Conditions for electroporation of mES cells with TagTK/NPT and clones generated

exp.	no. of cells	DNA	restriction enzyme	purification of DNA	voltage	capacity	incubation cells + DNA	no. of clones
1	10^7	8 µg	*PvuI*	P	240 V	475 µF	30 min on ice before pulse	9
2	10^7	33 µg	*SspI*	Q	240 V	475 µF	30 min on ice before pulse	1
3	10^7	25 µg	*SspI*	N	240 V	475 µF	30 min on ice before pulse	12
4	10^6	10 µg	*PvuI*	P	240 V	475 µF	30 min on ice before pulse	-
5	10^7	10 µg	*PvuI*	N	250 V	25 µF	30 min on ice before pulse	-
6	10^7	20 µg	*PvuI*	N	280 V	475 µF	15 min at RT after pulse	80

P = DNA purified with phenol
Q = purification of DNA via Qiagen columns
N = no purification, heat inactivation and precipitation by ethanol
RT = room temperature; exp.: experiment

In relation to the number of cells subjected to each electroporation very few G418 resistant clones could be established. As a positive control a vector expressing *npt* driven by the constitutively active human phosphoglycerate kinase (PGK) promoter was electroporated in

parallel with experiments 5 and 6. A significant number of clones could be generated in both experiments (54 and > 200 clones, respectively), showing that the applied electroporation and selection protocol as such is adequately efficient for the stable transduction of mES cells. The low number of clones obtained with pTagTK/NPT possibly indicates a lower transcriptional activity of the screening cassette.

2.2.2 Two cell clones show adequate and regulated luciferase expression and contain a single copy of pTagTK/NPT

To identify mES cell clones with regulated expression of the reporter cassette, the tagged mES cells were examined concerning their luciferase expression levels. Unexpectedly, very few clones expressed luciferase after addition of dox as shown in table 2. These clones were tested for luciferase activity in the presence and absence of dox (figure 9). All expressing clones were found to be regulated. Possible reasons for complete lack of reporter expression in the majority (i.e. ~ 94%) of G418 resistant clones will be discussed in chapter 3.2.1.

Table 2: Number of luciferase expressing, G418 resistant clones after induction with dox
G418 resistant clones were expanded, cultivated on gelatinized 6-wells with dox for 4 days and checked for luciferase activity. exp.: experiment

exp.	no. of luciferase expressing clones/ no. of G418 resistant clones
1	2/9
2	0/1
3	0/12
6	4/80

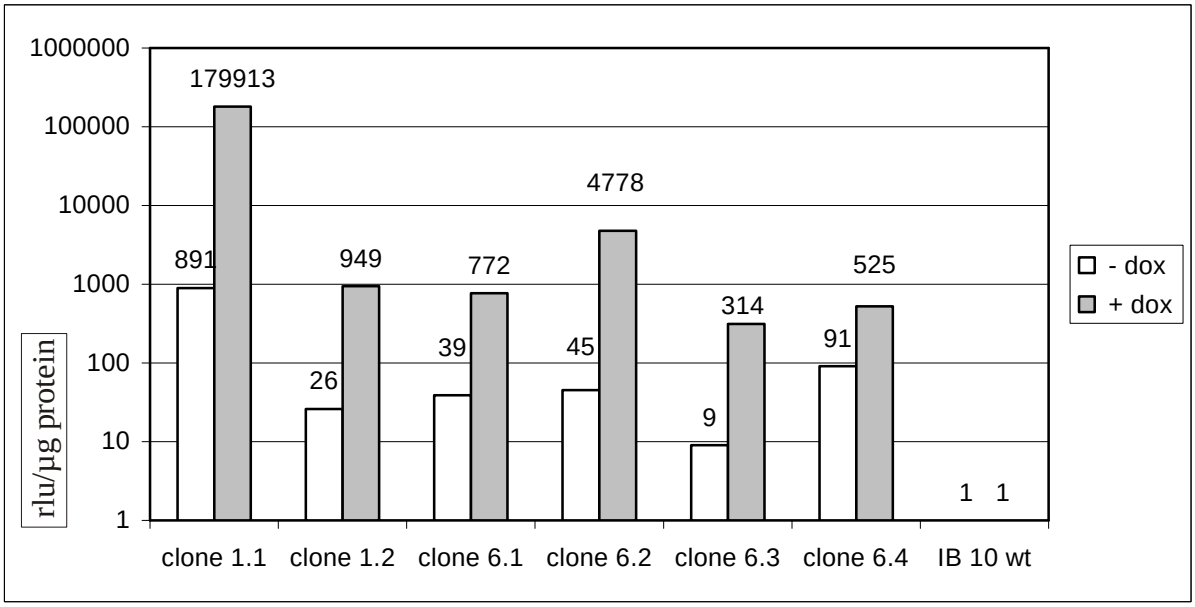


Figure 9: Regulated luciferase expression in mES cell clones tagged with pTagTK/NPT
Luciferase activity is given in relative light units (rlu)/µg protein; mES cells were cultivated on gelatinised 6-wells and dox was added for 4d for induced samples. Cells were harvested and checked for luciferase expression (details in "Material and Methods"). Clone 1.1 and clone 1.2 were obtained from experiment # 1; clones 6.1 - 6.4 from experiment # 6. IB10 wt cells were used as a negative control. Their luciferase activity did not exceed background levels and was hence set to 1. n = 1

Clones 1.1 and 6.2 displayed the highest luciferase expression and were therefore characterised for the number of integrated vector copies. Genomic DNA from each clone was digested with *EcoRV* which cuts within the cassette (figure 10). This enzyme also cuts in the genomic DNA next to the inserted vector, thus creating fragments whose length depend on the individual integration site. These fragments are detectable by hybridisation to appropriate probes, their number reflecting the number of individual integration sites. Sequences from the luciferase and *npt* genes were used as probes for a Southern blot. Both clones 1.1 and 6.2 displayed only a single band (figure 10). This was confirmed by additional experiments using different enzymes and probes (data not shown), proving that clones 1.1 and 6.2 carry a single integrated vector molecule.

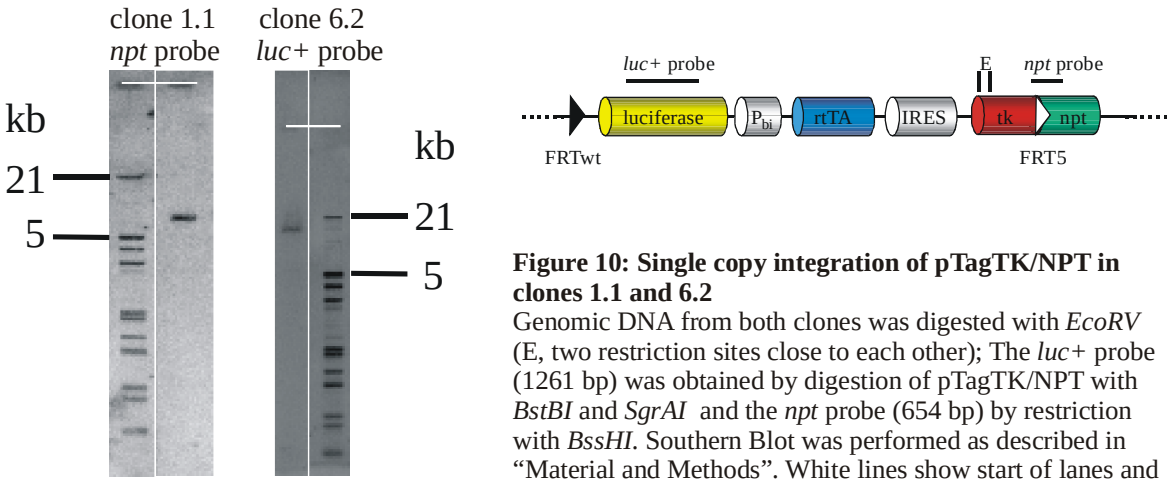


Figure 10: Single copy integration of pTagTK/NPT in clones 1.1 and 6.2
Genomic DNA from both clones was digested with *EcoRV* (E, two restriction sites close to each other); The *luc*+ probe (1261 bp) was obtained by digestion of pTagTK/NPT with *BstBI* and *SgrAI* and the *npt* probe (654 bp) by restriction with *BssHI*. Southern Blot was performed as described in "Material and Methods". White lines show start of lanes and single bands indicate that a single copy of the tagging vector has been integrated into the genomic DNA.

In order to more thoroughly characterise inducible expression, clones 1.1 and 6.2 were assayed for their luciferase activity by several independent experiments. Unexpectedly, luciferase activity strongly varied in different experiments for both clones although the cells had been treated identically. It will be shown later that this reflects an intrinsic property of the positive feedback cassette (chapter 2.2.2.2). For convenience, ranges of luciferase expression levels will be discussed and possible reasons for variations will be outlined in chapter 3.2.2.
In the following, clones 1.1 and 6.2 were characterised in detail concerning the time that is needed to reach complete induction and repression after addition or depletion of dox, respectively. Inducibility of the clones is measured by the regulation factor (expression levels in the induced state vs. basal expression), that gives information about how manifold expression of the transgene can be increased. In addition, autoregulated as well as constitutive luciferase expression from the ROSA26 locus was examined as a reference.

2.2.2.1 pTagTK/NPT supports fast induction and repression of autoregulated expression in clones 1.1 and 6.2

A prominent advantage of the tet system in contrast to other conditional gene expression techniques (e.g. induction of a gene by cre mediated removal of a stop cassette in front of the transcriptional start) is its reversibility. Gene expression in the tet on system is strictly dependent on the presence of the inducer and can be switched off again by depletion of the same. To determine the time required for induction/repression, luciferase levels of clones 1.1 and 6.2 were measured at different timepoints after addition/depletion of dox (switch on and switch off, respectively; figures 11 and 12).

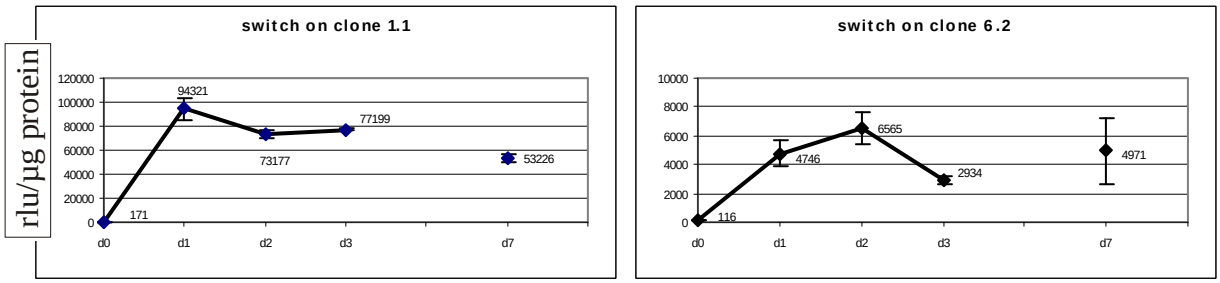


Figure 11: Switch on kinetics of luciferase expression for clones 1.1 and 6.2
Cell were cultured on gelatinised 6-wells. Samples were harvested at the indicated time points after addition of dox and luciferase levels were determined. rlu = relative light units; n = 2

As mentioned above, luciferase levels strongly vary in both clones. So, to determine the time the clones need to reach highest expression levels after induction, the first peak of each switch-on curve is considered as maximum.

When switched on clone 1.1 already reached the highest luciferase activity after 1d of induction. Clone 6.2 showed maximal expression after 2d when induced by dox. However, the measured luciferase level after 1d already amounted to 72% of the maximum and the values after 3 and 7 days were again lower than that after 2d.

Thus, in both clones induction is accomplished within 1 – 2d.

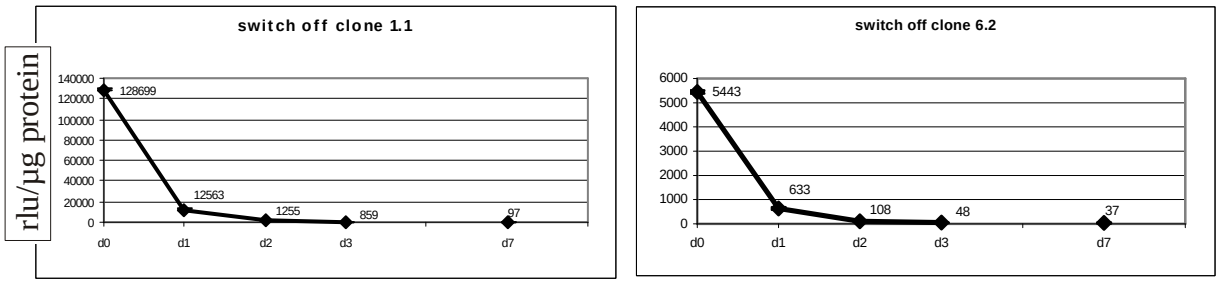


Figure 12 : Switch off kinetics of luciferase expression for clones 1.1 and 6.2
Cell were cultured in the presence of dox for 4d. They were then seeded to gelatinised 6-wells while dox was removed. Samples were harvested at the indicated time points after depletion of dox and luciferase levels were determined. rlu = relative light units; n = 2

In clone 1.1 luciferase expression was efficiently shut off after removal of dox (switch off). After 24 hours expression was already reduced to ~ 10% of the initial level and 2d later it was only slightly above basal expression. An almost identical induction/repression kinetic was observed for clone 6.2: luciferase activity was decreased to approximatly 12% after 24 h and basal expression level was reached after 2d. The fact that the shut down of expression is slower than induction of the system can be explained by the longer time that is needed for clearance of dox from the cells and for degradation of luciferase. These results show that the

tet on system is a useful tool for regulating gene expression that allows for a fast induction and suppression of the transgene.

2.2.2.2 Performance of clones 1.1 and 6.2 in comparison to expression from the ROSA26 locus

As indicated before, the autoregulated cassette in clones 1.1 and 6.2 is characterised by a significant fluctuation of expression. To exclude that this is a specific feature of the randomly obtained, unknown chromosomal integration sites, clones 1.1 and 6.2 were compared to ROSAautoLuc cells. In these cells a similar cassette had been integrated into the ROSA26 locus (figure 13), which is open in mES cells and in differentiated cells (Friedrich and Soriano, 1991; Zambrowicz *et al.*, 1997). As a control mES cells with luciferase constitutively expressed from the ROSA26 promoter were checked (termed ROSAEMLuc, figure 13). To evaluate the performance of the autoregulated cassette in the 3 different integration sites multiple experiments were performed to determine the level of luciferase expression and its fluctuations. Ranges of luciferase expressions, mean values and average regulation factors from multiple experiments are summed up in table 3.

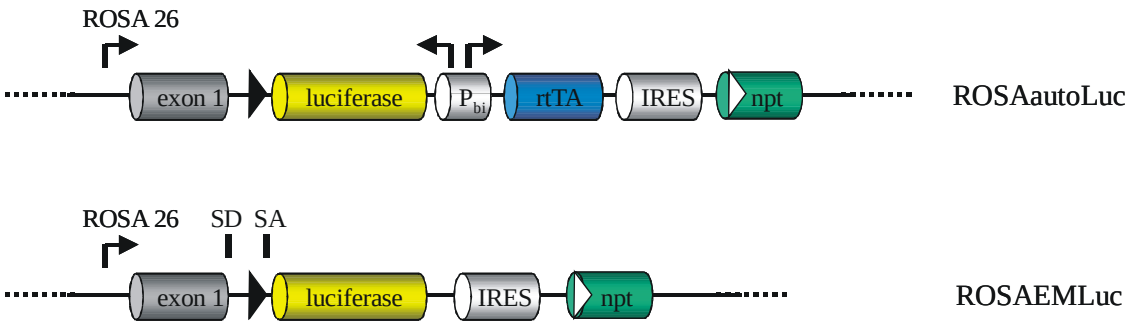


Figure 13: Schemes of an autoregulated and a constitutive cassette integrated in ROSA26
ROSAautoLuc ES cells carry a cassette similar to pTagTK/NPT which is integrated in the ubiquitously expressed ROSA26 locus. ROSAEMLuc cells in contrast express luciferase under the control of the constitutively active ROSA26 promoter. Both cell lines were generated by U. Sandhu and S. Bantner (described in detail in "Material and Methods") P_{bi} = tet dependent, bidirectional promoter; rtTA = reverse transactivator of the tet system, mutant rtTA2(S)-M2, IRES = Encephalomyocarditis virus internal ribosomal entry site, *npt* = neomycin phosphotransferase gene, SD = splice donor, SA = splice acceptor; black and white arrowheads: FRTwt and FRT5 sites, respectively

Table 3: Range of luciferase expression levels and regulation factors observed for clones 1.1, 6.2, ROSAautoLuc, and ROSAEMLuc
Values from kinetics experiments were included: induced = induced for 24 h or more, non induced = absence of dox for 72h or longer; ROSAEMLuc was only checked in the absence of dox; s.d. = standard deviations from average values, n = number of independent assays, RF = regulation factor, mean value induced state/mean value basal level

	uninduced [rlu/µg protein]				induced [rlu/µg protein]				average RF
	min	max	average	s.d.	min	max	average	s.d.	
clone 1.1	97	891	441 n = 20	78 %	2203	179913	47642 n = 34	94 %	108
clone 6.2	4	116	53 n = 20	66 %	175	6923	3347 n = 34	67 %	63
ROSAautoLuc	17	4144	1195 n = 5	144 %	381	14378	5263 n = 34	92 %	4
ROSAEMLuc (constitutive expression)	358761	889920	633836 n = 66	15 %	---	---	---	---	---

Standard deviations of mean values for autoregulated expression are extremely high (66% and more). In contrast, constitutive luciferase expression under control of the ROSA26 promoter deviates to 15%. Strong variations in luciferase activity are most likely associated to the design of the expression cassette, since they occur in three independent integration sites. Comparing mean luciferase levels in the induced state, clone 1.1 performs best (~ 9 times higher expression than ROSAautoLuc) while clone 6.2 shows half as much reporter activity as ROSAautoLuc. All three autoregulated clones display a certain basal activity, which amounts to 1% of average induced levels for clone 1.1, 1,6% for clone 6.2 and 22,7% for ROSAautoLuc. As a result of high basal activity ROSAautoLuc can apparently only be induced 4 fold. However, sample size for the uninduced state of ROSAautoLuc was comparatively small, so that this value might not be statistically relevant. Clone 6.2 has an average regulation factor of 63 and clone 1.1 can even be induced by two orders of magnitude. Switch on and switch off kinetics of ROSAautoLuc were comparable to that of clones 1.1. and 6.2 (data not shown).

On the whole, clone 6.2 shows less luciferase expression in the induced state than ROSAautoLuc, but its basal expression is much lower and its regulation factor consequently higher. In addition, mean values in absence/presence of dox indicate that the integration site of clone 1.1 is clearly superior to ROSA26 on mES cell level.

2.2.3 Evaluation of clones 1.1 and 6.2 in the differentiated state

Chromatin structures in mES cells are considered to be open as many genes are expressed in the pluripotent state. Transcription activity fundamentally changes during differentiation to mature cells and a multitude of loci are silenced. This affects not only endogenous genes but also randomly integrated transgenes. Accordingly, when establishing a transgenic mouse line there is a certain probability that the GOI is silenced completely or shows a heterogeneous expression pattern throughout the animal depending on the nature of its integration site(s). Due to unpredictable expression levels of the GOI after differentiation the standard protocol requires generation of more than one mouse line from ES cell clones with different integration sites.

To evaluate the expression characteristics of the tagged sites in clones 1.1 and 6.2 the next step of this work was to determine luciferase expression of both clones in the differentiated state. For this purpose several techniques were applied: (1) *in vitro* differentiation of the ES cells, (2) injection of ES cell clones into blastocysts and analysis of mouse embryonic fibroblasts (MEFs) from chimaeric embryos, (3) teratoma formation in immunosuppressed recipients and (4) analysis of transgenic mice themselves (chapters 2.4.1.2, 2.4.1.3).

2.2.3.1 Clones 1.1 and 6.2 maintain regulated luciferase expression after *in vitro* differentiation

mES cells were differentiated *in vitro* as described in “Material and Methods”. In short, the cells were kept in suspension culture in the absence of LIF for the formation of embryoid bodies (EBs). After 5 days EBs were collected and seeded to gelatinised culture dishes where they were allowed to further differentiate for 4 more days. The cells were then harvested, split and kept for 4d in the presence/absence of dox. This differentiation protocol does not direct the cells to specific lineages but should result in a pool of randomly differentiated progenitors of different cell types. Results of these experiments are given in figure 14. Switch on and switch off kinetics were not followed in this case because the number of cells obtained was not sufficient.

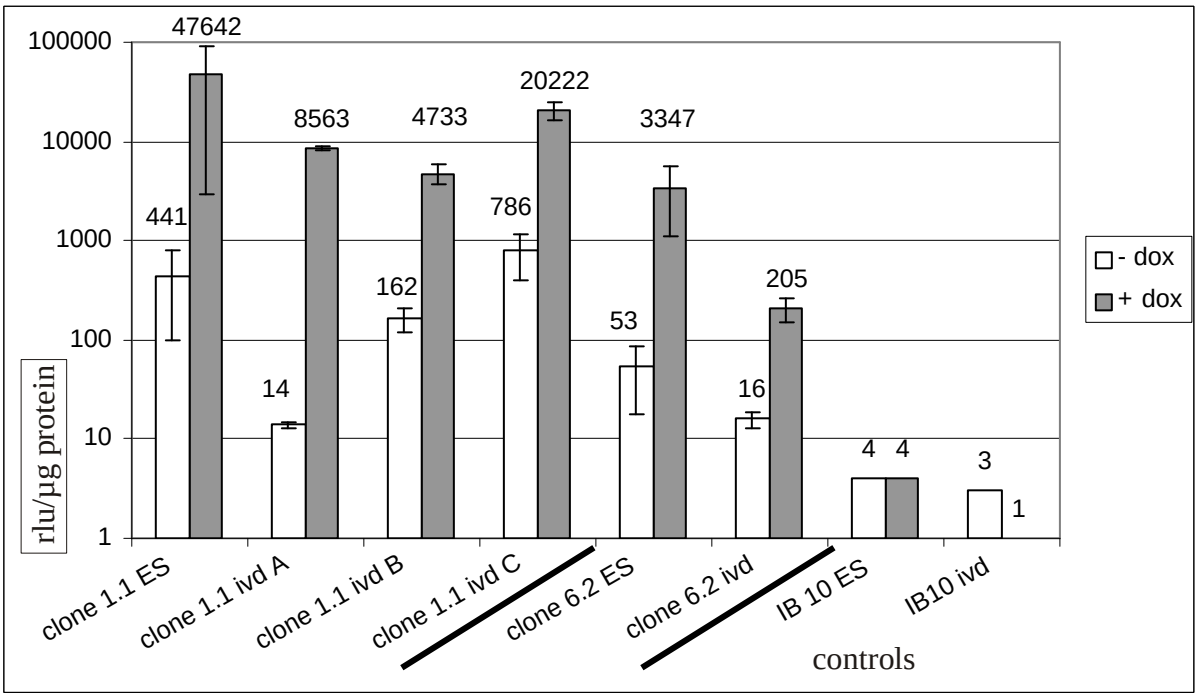


Figure 14: Luciferase expression levels of clones 1.1 and 6.2 before/ after *in vitro* differentiation
Differentiated cells were harvested after 4d in the presence/absence of dox. ES = mean values of ES cell state (sample size uninduced state: n = 20/induced state: n = 34 as given in table 3); ivd = *in vitro* differentiation; A, B, C = independent experiments (sample size uninduced/induced: A: 3/3, B: 6/6, C: 6/6, clone 6.2 ivd: 3/3); IB 10 wt cells were used as a negative control (ES cell state n = 6/6, ivd n = 6/6)

Luciferase expressions after differentiation of clones 1.1. and 6.2 lie within the ranges for the mES cell state indicated in table 3 (except uninduced state of *in vitro* differentiation experiment A with clone 1.1). Although values in the induced state are lower than the average of mES cells, it can be concluded that luciferase expression is not significantly silenced after *in vitro* differentiation. Also, regulation of expression is maintained. Variations between different experiments (experiments A-C with clone 1.1) may be due to the fluctuations of expression as observed for mES cells. Another possible cause might be changing contributions of different progenitor cell types to each differentiation cell pool.

2.2.3.2 Mouse embryonic fibroblasts (MEFs) isolated from chimaeric embryos derived from clone 6.2 show increased, regulated reporter gene expression

In order to gain more insight into the expression characteristics of clones 1.1 and 6.2 after differentiation, mES cells were injected into blastocysts and embryos were harvested at day 13,5. MEFs isolated were selected with G418 in the presence of dox to purify transgenic cells from chimaeric embryos. Resistant cells were subsequently analysed concerning luciferase

expression. Table 4 gives information about the total number of embryos isolated and the percentage of embryos with G418 resistant cells, i.e. chimaeric embryos.

Table 4: Analysis of injected embryos for transgenesis

	clone 1.1	clone 6.2
total number of embryos analysed	48	44
embryos yielding G418 resistant cells	-	10

No transgenic MEFs could be generated after injection of clone 1.1, which could have two reasons: either injection of mES into blastocysts cells was not efficient and no chimaeric embryos were present or the tagging cassette was not expressed in MEFs leading to loss of resistance to drug selection. The latter case would contradict the results obtained after *in vitro* differentiation and it cannot be ruled out at this point that ES cell injection of clone 1.1 into blastocysts completely failed. Since it was possible to generate chimaeric animals for establishing transgenic mice from this clone in a separate experiment (chapter 2.4.1.1), blastocyst injections of clone 1.1 to analyse chimaeric embryos were not repeated.
G418 resistant MEFs derived from clone 6.2 were analysed for their luciferase expression (figure 15).

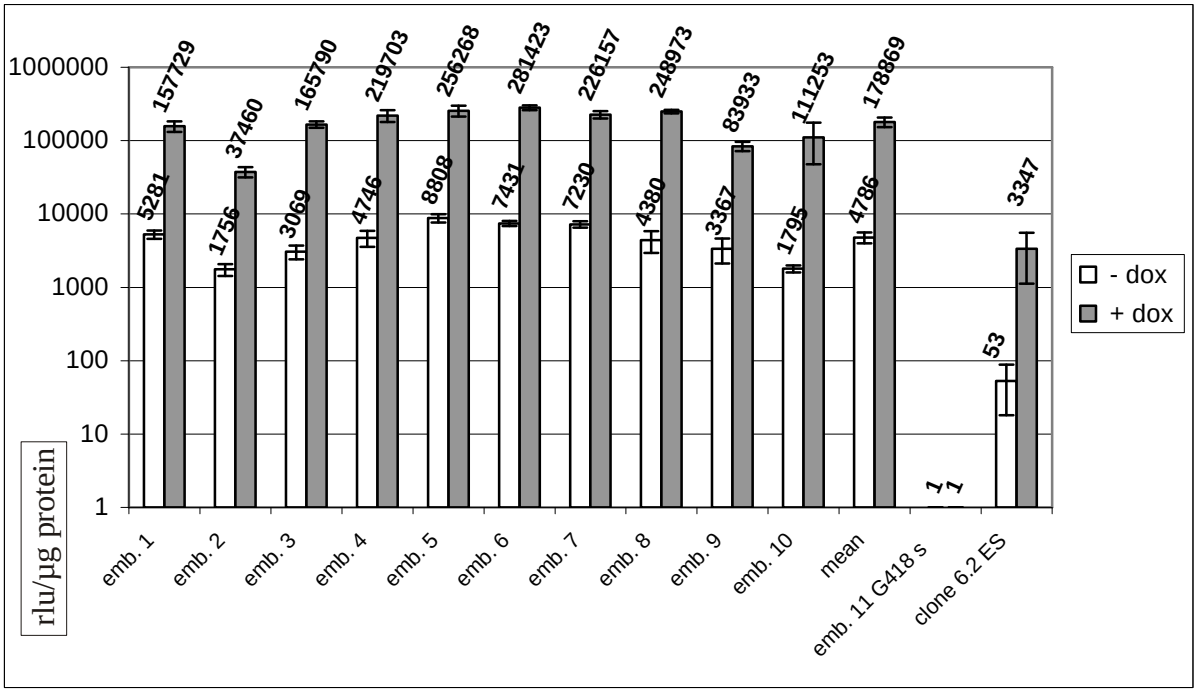


Figure 15: Luciferase expression in MEFs isolated from chimaeric embryos after injection of clone 6.2
Isolated MEFs were selected with G418 in the presence of dox for 10 days (controls dead), seeded to 6-wells and kept with/without dox for 6d; n = 3; emb. 11 G418 s = unselected MEFs from a G418 sensitive embryo; clone 6.2 ES = average values of mES cells, n = 20/34

In all cases expression was strongly increased in MEFs as compared to the mES cell state, the average of induced MEFs being more than 50fold higher than the mean readout observed for clone 6.2 mES cells. Also, basal expression levels increased and mostly even reached or exceeded mean induced state levels of mES cells. This results in a relatively small regulation factor that averages 37 which is lower than that observed for mES cells. On the whole, luciferase activities in MEFs isolated from different embryos were quite homogeneous and while overall expression levels were at least 15 times increased for the uninduced state and 5 times for the induced state (lowest MEF value/highest mES value), regulation was maintained to a significant level. Lack of strong variations of luciferase expression in MEFs could be explained by the fact that only a single experiment was performed. Still, expression levels of MEF pools from individual embryos were interestingly homogeneous.

2.2.3.3 Teratomas derived from clones 1.1 and 6.2 show luciferase expression

Another approach to compare performance of transgenic mES cells before and after differentiation is injecting them subcutaneously into isogenic or immunodeficient mice and examining the resulting tumors (i.e. teratocarcinomas, teratomas; e.g. described in Lange *et al*., 2000, Bonner *et al*., 2004). mES cells are pluripotent and randomly differentiate into cell types of all three germ layers in the "*in vivo*" environment of the recipient mouse.

Clone 1.1. and clone 6.2 ES cells were injected subcutaneously into immunodeficient Rag-2/ interleukin 2 receptor γ chain double knockout mice (generated as described in "Material and Methods"; model first described by Mazurier *et al*., 1999). As a negative control, IB10 wt cells were injected and mES cells constitutively expressing luciferase driven by the ROSA26 promoter served as a positive control (ROSAEMLuc). After 4 weeks teratomas had formed in some of the transplanted mice, but not in all (details and limitations concerning the technique of teratoma induction will be discussed in "Material and Methods"). Mice bearing tumors were then kept with/without dox administered via drinking water (2mg/ml as described by Kistner *et al*., 1996). Table 5 gives an overview about the animals that had developed tumors and the applied induction protocol. Luciferase intensity emitted from the teratomas was determined by *in vivo* bioluminescent imaging (BLI) with a Xenogen IVIS 200 system. Briefly, mice were anaesthetised, depilated and after i.p. injection of luciferin emission of light was measured. Individual tumors grew with different growth rates, so that only qualitative analysis of luciferase expression after differentiation was possible.

Table 5: Overview of mice bearing teratomas and treatment with dox
Animal #2 had to be killed due to tumor size.

mouse no.	left side	right side	administration of dox
1	-	IB10 wt	8 d –dox ; 4 d + dox
2	-	ROSAEMLuc	8 d –dox; killed
3	clone 1.1	-	8 d +dox ; 4 d –dox
4	clone 1.1	-	8 d –dox ; 4 d + dox
5	clone 1.1	ROSAEMLuc	8 d +dox ; 3 d –dox
6	clone 6.2	ROSAEMLuc	8 d -dox ; 3 d +dox

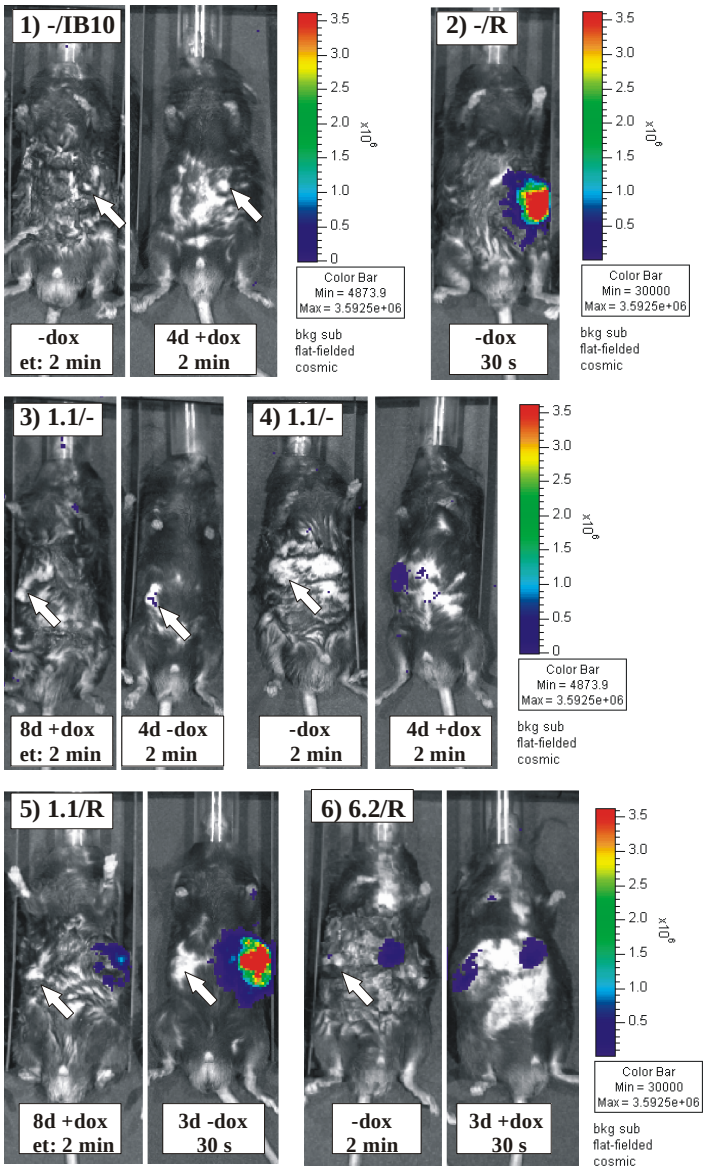


Figure 16: In vivo bioluminescent imaging of mice bearing teratoma
Mice were administered 2mg/ml dox in the drinking water as indicated in table 5. For in vivo bioluminescent imaging they were anaesthetised with 2-2,5% isoflurane (Isoflo, Abbott), treated with depilatory cream and injected i.p. with 100µl of luciferin (30mg/ml in PBS; Synchem OHG). They were put into the Xenogen IVIS 200 imaging chamber and luciferase activity was measured under constant inflow of isoflurane via nose cones. Rainbow scales indicate number of photons measured per second x cm2 x steradian. Imaging conditions: field of view: 19.6; binning: medium, aperture: f1, et = exposure time, given for each image. Exposition time had to be reduced from 2 min to 30 s when signal intensity was too strong. White arrows indicate positions of tumors when these do not emit any signal. Numbers of panels and animals are identical and the source of teratomas is indicated as follows: IB10 = tumor derived from IB10 wt, R = tumor derived from ROSAEMLuc, 1.1 = tumor derived from clone 1.1, 6.2 tumor derived from clone 6.2, - = no tumor

Controls

The IB10 tumor (mouse 1) did not give any signal both in the absence and presence of dox as expected (figure 16 panel 1) Mice injected with ROSAEMLuc cells showed very high luciferase activity from these teratomas (figure 16 panels 2, 5 and 6, right side). Differences in signal intensity were due to varying tumor sizes.

Clone 1.1

The three tumors derived from clone 1.1 behaved inconsistently when assayed for luciferase expression. Mouse 4 showed no basal expression and could be induced by application of dox (figure 16 panel 4). Unexpectedly, animal 3 only gave a very faint signal after removal of the inducer (figure 16 panel 3). Mouse 5 did not display any luciferase expression from the clone 1.1 tumor (figure 16 panel 5). However, this might be due to the very intensive signal from the ROSAEMLuc control tumor which could decrase sensitivity so that a potential weak signal from the clone 1.1 teratoma might not be detectable (note that exposition time had to be decreased from 2 min to 30 s for the second image of panel 5).

Clone 6.2

Without dox, no basal expression could be observed for the clone 6.2 teratoma of mouse 6 (figure 16 panel 6). After three days of dox administration, a clear signal could be deteced (note decrease of exposition time to 30s instead of 2min), indicating that autoregulated expression is well inducible after differentiation of cells.

Both clones 1.1 and 6.2 show expression of luciferase after differentiation of mES cells by teratoma formation. Regulation is maintained in the case of clone 6.2 as judged by mouse 6. Expression of differentiated clone 1.1 cells seems to be heterogeneous and not reliably regulated. A possible reason for this could be restriction of expression to a certain tissue or cell type due to the chromosomal integration site of the cassette. Varying contributions of differentiated lineages in each tumor would then lead to heterogeneous expression. The results obtained after teratoma formation will be discussed later in context with the data obtained from transgenic mice (chapters 3.2.3, 3.3.1).

Taking together the results from all differentiation approaches it can be concluded that both clone 1.1 and 6.2 maintained luciferase expression after *in vitro* and *in vivo* (MEFs, teratoma) differentiation. While clone 6.2 was still well regulated by addition/depletion of dox in all

performed assays, regulation of clone 1.1 was accomplished *in vitro* whereas it was ambiguous in teratomas.

2.3 Targeting mES cells via Flp mediated recombination

The applied strategy was designed to specifically reuse a defined locus that is tagged by recombination target sites. While expression of randomly integrated transgenes is heterogeneous and strongly depends on their integration sites, this method can provide predictable expression characteristics of the site specifically integrated GOI. To show that the tagged loci of clones 1.1 and 6.2 can be reused via RMCE, targeting of both cell lines was followed.

Cells with stably integrated pTagTK/NPT are resistant to G418 and sensitive to GCV in the presence of dox. For exchange of the reporter cassette a targeting vector and a Flp recombinase expression plasmid are cotransfected into the cells. Flp recombinase mediates recombination between identical FRT sites so that the FRT wild type and mutant 5 sites from the tagged locus will only interact with their respective counterparts in the targeting vector. The incoming cassette eliminates the *tk* reading frame and provides a functional start codon for the truncated *npt*, leaving targeted clones resistant to both GCV and G418. Accordingly, cells are selected with both drugs after cotransfection of targeting and Flp recombinase vectors (figure 17).

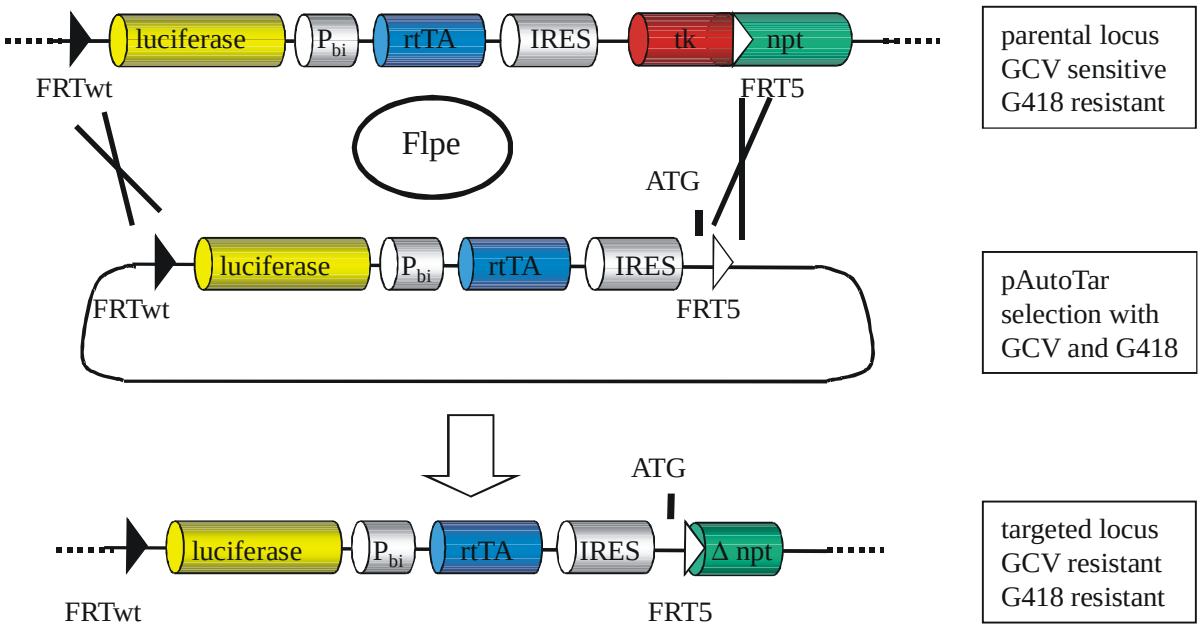


Figure 17: pTagTK/NPT cassette exchange strategy
Flpe expression and targeting plasmids are cotransduced into cells with the FRT tagged locus. Flpe mediates recombination between homotypic sites thus exchanging the cassettes. Selection with GCV and G418 in presence of dox recovers only correctly targeted clones.
P_{bi}: bidirectional, tet dependent promoter; rtTA: reverse transactivator variant rtTA2(S)-M2; *tk/npt*: fusion of the thymidine kinase gene with the neomycin phosphotransferase gene; IRES: Encephalomyocarditis virus internal ribosombal entry site; FRTwt and FRT5 sites are indicated by black and white arrowheads, repectively.

Sustained G418 pressure forces the cells to consistently express *npt*. This should exclude resistance by simple inactivation of *tk*. To evaluate the necessity of retained G418 pressure in addition to GCV selection, a growth assay under different selection conditions with clone 1.1 and IB10 wt cells was performed (data not shown). The results indeed showed that eradication of dox induced pTagTK/NPT bearing cells by GCV was accomplished in the presence of G418 but insufficient in its absence. IB10 wt cells were not affected by the GCV concentrations applied as expected.

Thus, for later targeting experiments cells were selected in presence of all three drugs since this regimen reliably kills all cells carrying the parental construct. Clone 6.2 was treated in the same way as it was expected to behave identical to clone 1.1 with respect to its selection properties.

2.3.1 Attempts to target clones 1.1 and 6.2

Targeting, i.e. specific exchange of the FRT flanked sequence for a new cassette, is routinely performed by cotransfection of the targeting plasmid itself along with a Flp recombinase expression vector (e.g. Seibler *et al.*, 1997). Although various transfection methods, selection

regimens and targeting vectors were applied in a multitude of experiments, it was impossible to target either clone 1.1 or clone 6.2.

Efficiency of RMCE is influenced by multiple parameters that include a) integrity of FRT sites, b) transfection efficiency, c) design of the targeting plasmids, d) adequate Flpe expression levels and e) appropriate selection protocols. These issues will be discussed in the following paragraphs.

a) Integrity of FRT sites is essential for efficient cassette exchange. Sequencing of FRT sites in pTagTK/NPT plasmid DNA revealed that both wild type and mutant sequences are intact. To exclude the possibility that FRT sites were mutated or lost during integration into chromosomal DNA, genomic DNA of clone 6.2 was used to amplify the regions surrounding these sites (data not shown). Resulting PCR fragments were purified and sequenced with appropriate primers. This confirmed that also in clone 6.2 both FRT sites are unimpaired. FRT sites of clone 1.1 were not examined.

b) For cotransduction of targeting and Flpe expression plasmids various electroporation and transfection protocols were applied (for electroporation view e.g. Liu *et al*., 2006b; Cobellis *et al*., 2005; transfection reagents applied: FuGENE6, Roche; Lipofectamine2000, Invitrogen). These protocols had been successfully used to target the ROSA26 locus by RMCE, albeit in combination with a different selection strategy that will be delineated later. Although clones 1.1 and 6.2 were transfected under identical conditons as the ROSA26 FRT tagged clone, targeting was not succesful. As a definite proof that DNA was properly delivered to the cells, eGFP and DsRed expression vectors were cotransfected using Lipofectamine 2000 and transfection efficiency was determined by FACS analysis. Both plasmids are sufficiently expressed in transfected cells and more than 15% of all cells had taken up both vectors (data not shown). Consequently, transfection efficiency is not the limiting step for targeting clones 1.1 and 6.2.

c) Initially, the vector pAutoTar was used for targeting attempts (figure 18), which is composed of the same features as the tagging vector except that the *tk* gene is missing and a startcodon is included to conserve *npt* expression. Targeting with pAutoTar would hence allow to directly compare luciferase expression levels and regulation before and after cassette exchange. However, as first targeting experiments did not yield any clones the strategy was changed to transfecting simpler constructs that are not dox dependent but constitutively expressed. When targeting attempts with the

plasmid pFCL3ETKNF5 did not work out, a non site specific selection approach was applied using the plasmid pFCL3EPF5 which renders targeted cells resistant to puromycin. Of course, also random integration of the cassette will result in resistant clones. Cobellis *et al.* (2005) showed that 2 out of 30 clones were correctly targeted using non site specific selection after RMCE while Masui *et al*. (2005) even observed an efficiency of 50%. In this case, random integration scarcely took place: only 6 puromycin resistant subclones of clone 1.1 and 7 of clone 6.2 could be generated from several different experiments, using $3{,}3x10^6$ and $3{,}5x10^6$ cells on the whole. After analysis by PCR or Southern blot all these clones turned out to be established solely by random integration of the targeting vector but not from correct targeting events (data not shown).

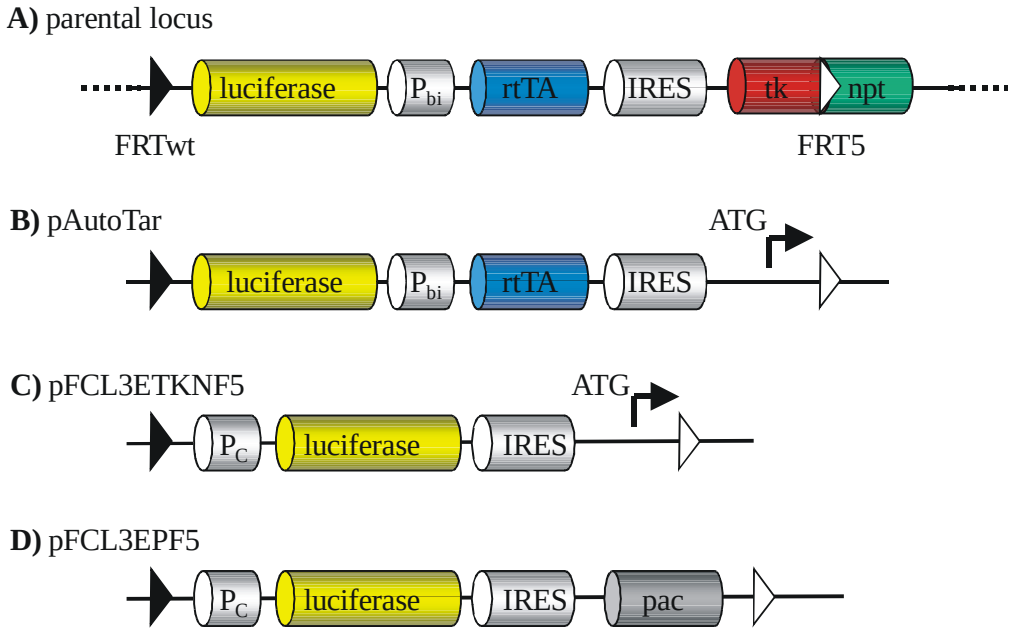


Figure 18: Schematic view of pTagTK/NPT tagged locus and targeting vectors
A) Schematic view of the tagged locus B) Autoregulated, tet dependent targeting cassette C) Constitutively active targeting cassette for site specific selection via loss of *tk* D) Constitutively active targeting cassette for non site specific selection by puromycin; black and white arrowheads: FRTwt and FRT5 sites, respectively; P_{bi}: bidirectional, tet dependent promoter; rtTA: reverse transactivator variant rtTA2(S)-M2; *tk/npt*: fusion of the thymidine kinase gene with the neomycin phosphotransferase gene; IRES: Encephalomyocarditis virus internal ribosombal entry site; P_C: CAGGS promoter (Araki *et al*., 1997); *pac*: puromycin acetyl transferase gene

d) Efficient targeting is dependent on the activity of the Flpe recombinase. To drive Flpe expression the CAGGS promoter (Araki *et al*., 1997) was used, which works well in particular for mES cells (Schaft *et al*., 2001). To enforce Flpe expression the so called "puro shock" method was tested: the Flpe vector contains an IRES - puromycinacetyl

transferase (*pac*) element. After transfection, puromycin is added for two days which forces the cells to transiently express *pac* and also *Flpe*. As transient Flpe expression did not yield targeting, clones 1.1 and 6.2 were transfected with the above described Flpe-puro vector and selected with puromycin. 4 stably Flpe expressing subclones were randomly chosen for each clone 1.1 and 6.2 and transfected with the targeting vector. However, targeting also failed when Flpe was constantly present.

In principle, FRTwt and FRT5 sites should not interact as they are not identical. If there still was a certain level of recombination between these sites, excision of the sequence flanked by them could occur leading to loss of G418 resistance. To exclude this, a Southern blot was performed showing that for all 8 subclones the cassette was kept intact even if Flpe was stably expressed (data not shown).

2.3.1.1 TK/NPT selection strategy for targeting

Site specific selection was performed by induction with dox 24h after transfection and addition of GCV and G418 at least another day later to ensure sufficient *tk* and *npt* expression levels. Resistance to GCV is dependent on the absence of TK protein; hence, complete degradation of the protein in targeted subclones is a prerequisite for successful selection. Taking this into account selection with GCV was started 4 to 6d after transfection in some experiments.

Another critical factor is the so called “bystander effect”, i.e. non targeted cells in close proximity to targeted subclones lead to cell death of these correctly exchanged cells. To avoid this, cells were seeded sparsely on dishes with feeder cells prior to selection.

Transfection for targeting was performed both with and without dox in several experiments. On the one hand, induction by dox and transcription should open the chromatin structure and might make it more accessible for Flpe and the targeting plasmid. On the other hand, the transcriptional apparatus and the whole process of transcription might block the FRT tagged site for targeting. Both approaches were not successful, though.

Since all the potential causes of defect described above could be excluded, lack of targeting in clones 1.1 and 6.2 could be either due to integration sites which are refractory to RMCE or to an extremely low efficiency of the applied exchange strategy in mES cells. The fact that both clones readily express luciferase upon induction indicates that the tagged loci are principally accessible for modifications and hence dismisses the first possibility. As *pac* requires higher expression levels to mediate resistance to puromycin than *npt* to detoxify G418, relatively low

expression from both loci might be the cause for failure of the non site specific selection exchange strategy as described by Cobellis *et al*. (2005). Accordingly, an alternative tag and target strategy described below was pursued.

2.3.2 Efficient targeting of ES cell clones with the PAC/ΔNPT selection strategy

It was assumed that the design of the selection strategy contributed to the failure of targeting clones 1.1 and 6.2. Thus, an alternative selection strategy that had succeeded in targeting the ROSA26 locus was employed for tagging and targeting of random loci. This strategy involves selection for puromycin resistance in the tagging step and site specific complementation of an ATG deficient *npt* gene after targeting. Hence, a new tagging construct (pTagPAC/ΔNPT) was developed and used for tagging IB10 mES cells (new tag and target strategy depicted in (figure 19).

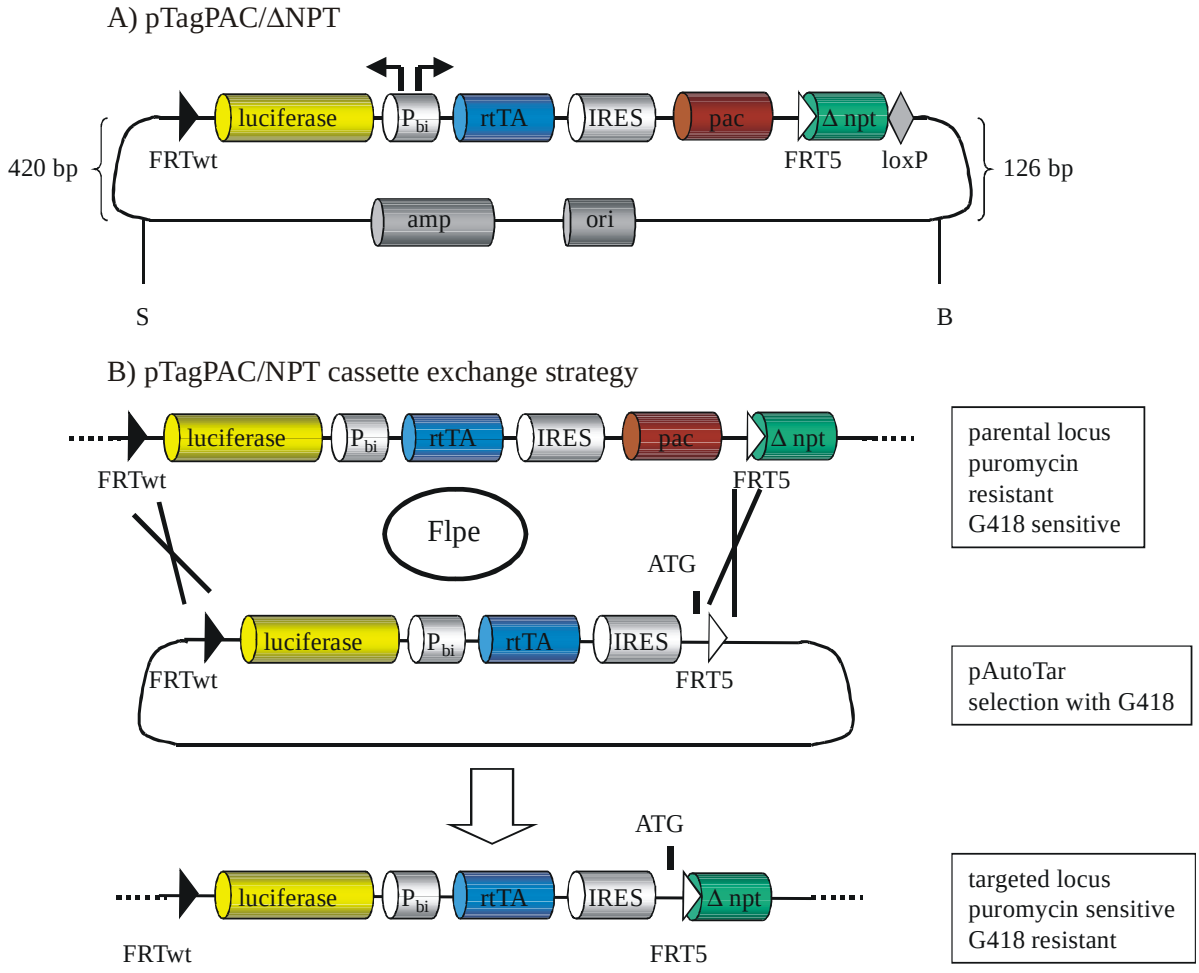


Figure 19: new tag and target strategy with pTagPAC/ΔNPT
A) Scheme of pTagPAC/ΔNPT which mediates puromycin resistance in the presence of dox; ΔNPT is not functional. B) Targeting of the parental locus renders cells G418 resistant and puromycin sensitive. P_{bi}: bidirectional, tet dependent promoter; rtTA: reverse transactivator variant rtTA2(S)-M2; pac: puromycin acetyltransferase gene; Δ*npt*: truncated neomycin phosphotransferase gene; IRES: Encephalomyocarditis virus internal ribosombal entry site; FRTwt and FRT5 sites are indicated by black and white arrowheads, repectively.

2.3.2.1 Generation of ES cell clones tagged with pTagPAC/ΔNPT

pTagPAC/ΔNPT plasmid DNA was linearised following two different approaches: digestion with *BstZ17I* alone linearised the vector in the backbone while double digestion using *BstZ17I* and *SspI* removed a large part of potentially unfavourable bacterial sequences (figure 19). Linearised DNA was electroporated into IB10 mES cells as indicated in table 6.

Table 6: Conditions for electroporation of mES cells with pTagPAC/ΔNPT and clones generated
Experiments 3 and 4 were performed with the Microporator technology from Peqlab using 100µl electroporation tips. Cells were cultivated with dox 24h after electroporation and 1 µg/ml puromycin was added to the media another day later. Clones were picked after 14 and 16d of selection (plates I + II and plate III, respectively).

exp.	no. of cells	DNA	restriction enzyme	purification of DNA	voltage	capacity	incubation cells + DNA	no. of clones
1	$5x10^6$	45µg	*BstZ17I*	N	280 V	475 µF	15 min at RT after pulse	-
2	$5x10^6$	45µg	*BstZ17I* + *SspI*	N	280 V	475 µF	15 min at RT after pulse	-

exp.	no. of cells	DNA	restriction enzyme	purification of DNA	voltage	pulse width	pulse number	no. of clones
3	$5x10^6$	100 µg	*BstZ17I*	P	1350 V	20 ms	2	44
4	$5x10^6$	100 µg	*BstZ17I* + *SspI*	P	1350 V	20 ms	2	157

P= DNA purified with phenol
N = no purification, heat inactivation and precipitation by ethanol
RT = room temperature
ms = milliseconds

Numbers of resistant clones could be significantly increased by using an adapted electroporation protocol for the transduction of pTagPAC/ΔNPT that probably causes less cell death. Thus, a larger amount of clones was available for screenings.

To identify luciferase expressing clones, puromycin resistant cells were grown with dox on gelatinised microtiter plates to confluence. These plates were then imaged with the Xenogen IVIS 200 system as depicted in figure 20.

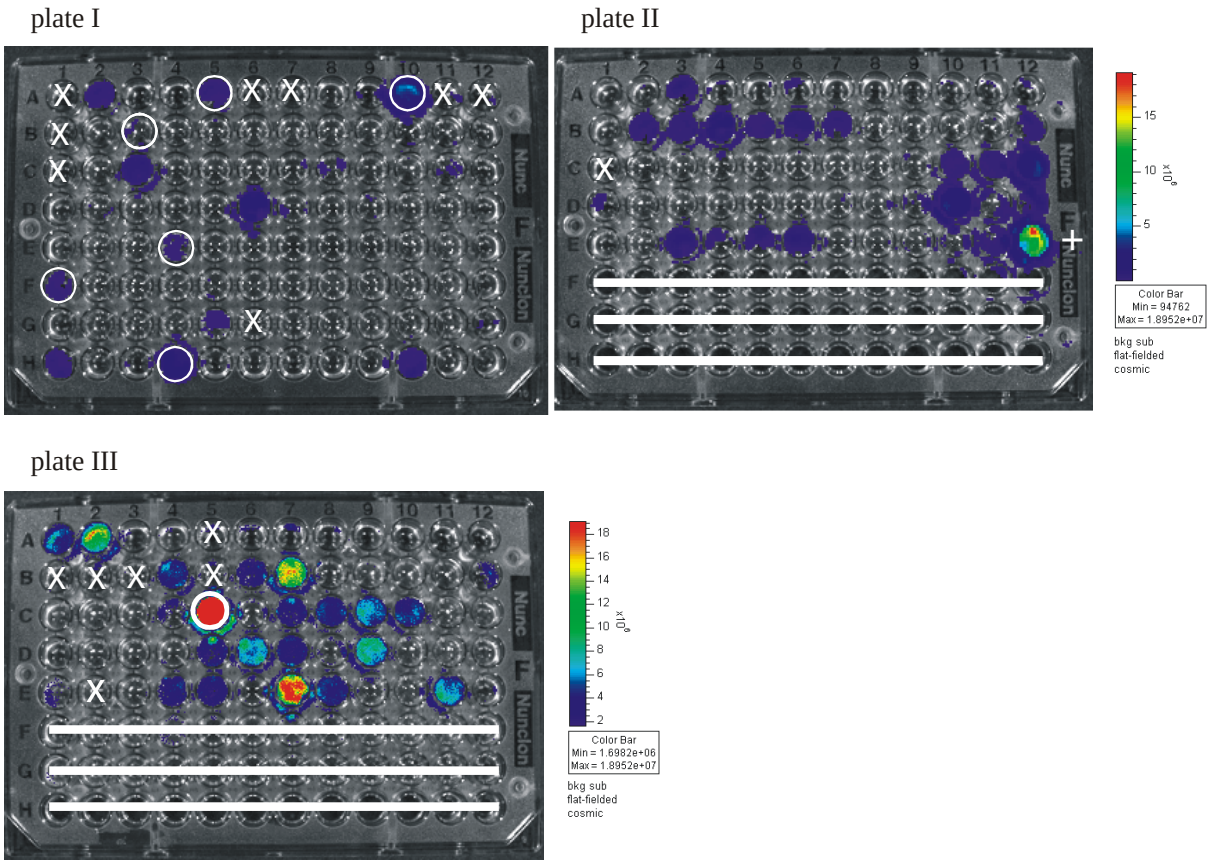


Figure 20: Evaluation of luciferase activity in pTagPAC/ΔNPT tagged clones
Clones were grown on gelatinised microtiter plates and induced with dox for 3 days. For bioluminescence imaging the medium was removed, cells were washed once with PBS and 50µl of luciferin/ATP reaction buffer were added to each well (described in detail in "Material and Methods"). Plates were imaged under the following conditions: field of view: 13; binning: medium for A and B, small for C; aperture: f1, exposure time: 20 s for A and B, 1 s for C. Empty wells are marked by an "x" or a white line. As a positive control, the ROSAEMluc clone constitutively expressing luciferase under control of the ROSA26 promoter was cocultured (indicated by +). Clones that were selected for targeting experiments are designated by a white circle. Clones carrying pTpagPAC/ΔNPT digested with *BstZ17I*: plate I A1-C12, plate III A1 – B6; clones tagged with double digested pTagPAC/ΔNPT (*BstZ17I* + *SspI*): plate I D1 – H12, plate II A1 – E11, plate III B7 – E12 (~20 small clones from this electroporation were not picked).

In line with the results from clones tagged with pTagTK/NPT, most pTagPAC/ΔNPT clones again did not express luciferase although they were resistant to the respective selection drug in the presence of dox. However, while only 6% (6/102) of pTagTK/NPT tagged clones showed luciferase activity, approximately 20% of the 201 pTagPAC/ΔNPT clones gave a clear signal. Possible reasons for the relatively low percentage of luciferase expressing clones among drug resistant cells will be discussed in detail in chapter 3.2.1.

To evaluate the new RMCE selection strategy, 7 inducible pTagPAC/ΔNPT tagged clones were chosen as indicated in figure 20 for targeting experiments. The number of integrated copies was not determined beforehand because of time constraints.

2.3.2.2 Successful targeting of 5 independent pTagPAC/ΔNPT tagged clones

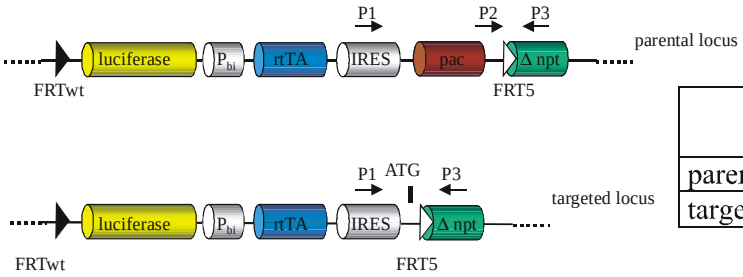
Clones IA5, IA10, IB3, IE4, IF1, IH4 and IIIC5 were cotransfected with Flpe expression and targeting plasmids. Transfection conditions and numbers of G418 resistant clones obtained after selection are given in table 7. The targeting plasmid pAutoTar had the same autoregulated arrangement of components as the tagging vector so that expression levels before and after targeting were directly comparable (figure 19).

Table 7: Transfection of pTagPAC/ΔNPT tagged clones for targeting and number clones obtained after selection
$5x10^5$ cells of each clone were seeded to gelatinised 6-wells and cotransfected with 2,5 µg pCflpe and 1,5 µg pAutoTar using 7 µl of Lipofectamine 2000 (Invitrogen) transfection reagent as recommended by the manufacturer. Cells were transferred to feeder coated 10cm dishes and dox was added 24 h after transfection. 0,4 mg/ml G418 were added another 24h later. Colonies were counted and if possible 4 subclones each were picked after 12 days of selection.

parental clone	number of $G418^r$ subclones
IA5	~ 440
IA10	-
IB3	112
IE4	-
IF1	~ 230
IH4	~ 140
IIIC5	~ 270

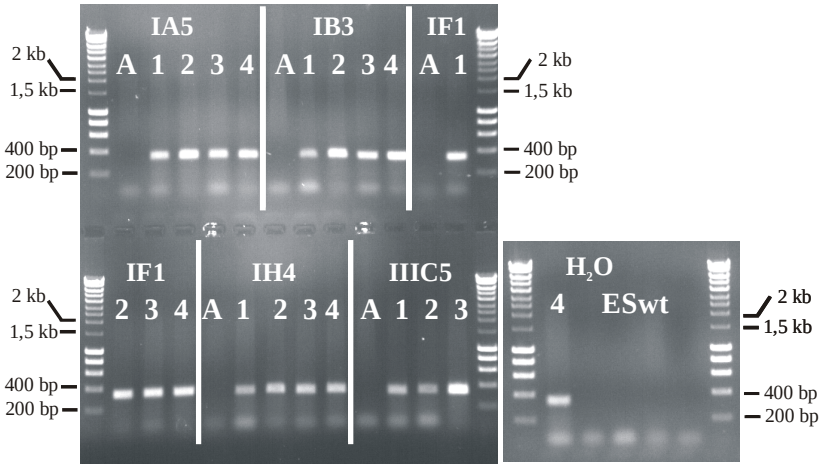
All tagged clones except IA10 and IE4 yielded significant numbers of G418 resistant subclones. 4 subclones each were analysed for correct targeting by PCR as shown in figure 21.

1) Primer binding sites and expected fragment sizes



	PCR A P1 + P3	PCR B P2 + P3
parental locus	1,8 kb	820 bp
targeted locus	300 bp	-

2) PCR A with primers P1 and P3



3) PCR A with primers P1 and P3; PCR B with primers P2 and P3

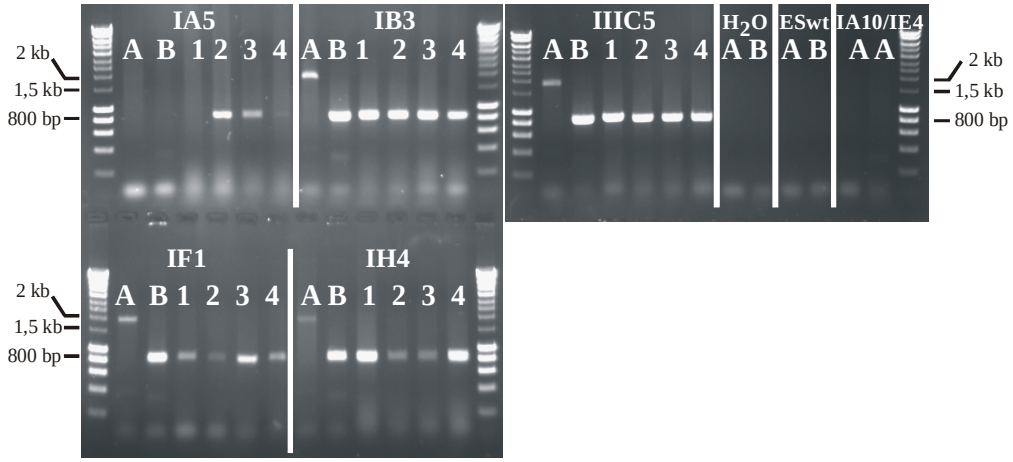


Figure 21: PCR analysis of G418 resistant subclones

Cells were cultured on gelatinised 96-wells, genomic DNA was harvested and PCR performed with 3µl of this template as described in "Material and Methods".

1) Schemes of parental and targeted locus with primer binding sites (not true to scale); the table gives information about expected fragment sizes for each primer combination and locus configuration. 2) PCR A with parental clones (indicated with "A") and 4 $G418^r$ subclones. The 300bp band in subclones shows that targeting has taken place. PCR failed for parental clones due to suboptimal amplification con ditions for the expected 1,8kb fragment. The last two lanes are empty. 3) PCR A for parental clones ("A") and PCR B for parental clones ("B") and subclones. The 1,8kb band could sucessfully be generated in PCR A this time (except for clone IA5 and clones IA10and IE4). PCR B reveals the expected 820 bp band for the tagged constellation in the parental clones as well as in the targeted subclones. PCR B did not yield any band for IA5 and its subclones 1 and 2. However, as this band is clearly detectable in clones 3 and 4 it can be concluded that the parental construct is present in IA5, 1 and 2.

P_{bi}: bidirectional, tet dependent promoter; rtTA: reverse transactivator of the tet system rtTA2(S)-M2; IRES: Encephalomyocarditis virus internal ribosombal entry site; *pac*: puromycin acetyltransferase gene, *npt*: neomycin phosphotransferase gene; ESwt: IB10 wt genomic DNA

PCR analysis using a primer pair binding in the IRES element and in *npt* (PCR A) clearly showed that all 4 subclones of each parental clone have been successfully targeted (figure 21 panel 2). PCR specifically detecting the parental configuration (PCR B) also revealed the presence of non targeted cassette(s), thereby implying that clones IA5, IB3, IF1, IH4 and IIIC5 carry more than one copy of pTagPAC/ΔNPT. Multiple copies of the tagging cassette may have led to the extremely high numbers of targeted subclones, since only one copy needed to be exchanged to gain resistance to G418. Nevertheless, the PAC/ΔNPT selection strategy is also efficient for a single tagged locus (60-80% correctly targeted/resistant clones when targeting ROSA26 by RMCE; personal communication from S. Bantner).

Lack of targeting in clones IA10 and IE4 could be due to loss of part of the *npt* gene during integration of the tagging cassette into the chromosomal DNA, since PCR B did not yield a band for either clone (figure 21 panel C).

Altogether, these results show that the PAC/ΔNPT selection strategy works efficiently to retarget randomly tagged loci in mES cells via RMCE.

2.4 Potential applications for targetable mES cells supporting autoregulated, tet dependent expression

The work described above yielded achievements that give direction to future experiments. It could be shown that the PAC/ΔNPT selection strategy is well suited for tagging/targeting approaches mES cells. In contrast to the strategy employing TK as a negative marker previously designed to select for targeted subclones, complementation of a truncated *npt* gene by the incoming vector efficiently allows to reuse tagged sites of interest. Accordingly, screening of pPAC/ΔNPT tagged cells will be performed to identify clones that qualify for application by a) carrying a single integrate b) displaying good inducibility by dox and c) retaining expression and regulation in differentiation assays.

Potential applications of mES cell clones carrying RMCE targetable loci which support regulated gene expression include rapid generation of transgenic mice as well as *in vitro* differentiation.

Differentiation of mES cells towards cell types of specific lineages *in vitro* represents a source for somatic cells that circumvents the time consuming process needed to establish transgenic mice. Transgenes might be specifically integrated into precharacterised loci by RMCE, thus ensuring predictable and - if desired - regulated expression levels. As proof of principle, clone

6.2 and an mES cell clone carrying a tet regulated reporter cassette in ROSA26 were differentiated to hematopoietic progenitors as described by Pilat *et al*. (2005; chapter 4.7.11). Finally, transgenic mice were established from clones 1.1 and 6.2 and analysed by bioluminescent imaging for validation of the differentiation assays used to characterise clones 1.1 and 6.2 and to evaluate performance of the autoregulated cassette *in vivo*.

2.4.1 Evaluation of autoregulated cassettes in transgenic mice

2.4.1.1 Generation of transgenic mice

Differentiation experiments indicated that luciferase expression of clones 1.1 and 6.2 is maintained after commitment of mES cells and can still be regulated by dox at least in the case of clone 6.2. To characterise rerporter gene activity concerning potential tissue specificity and induction levels *in vivo*, transgenic mice were established from both clones in the central mouse facility under guidance of the "Experimental Immunology" department headed by Dr. Werner Müller. Briefly, ES cells were injected into Balb/C blastocysts which were then transplanted into foster animals. Resulting chimaeras were identified by coat colour (silver grey for IB10 derived fur on white Balb/C background). Male animals showing the highest contribution of injected ES cells as judged by coat colour chimaerism were chosen for matings with Balb/C females. Their offspring was genotyped by analysis of tail tip DNA either by Southern blot or PCR. Table 8 shows an overview of generated chimaeras. Both clones yielded germline transmission and transgenic mice were bred for evaluation experiments.

Table 8: Chimaeras obtained after blastocyst injection for generation of transgenic mice
m = male; f = female; only male chimaeras were mated because IB10 derived ES cells are male. Female chimaeras consequently very rarely yield transgenic offspring. One of the clone 1.1 derived chimaeras died 5 weeks after birth. However, this is most probably not related to chimaerism but due to general effects during blastocyst injection, as 3 non chimaeric animals died as well.

injected clone	chimaeras	transgene transmitting/ mated males
clone 1.1	2 m (one died before mating)	1/1
clone 6.2	9 m / 2f	3/5

2.4.1.2 Luciferase expression in mice derived from clones 1.1 and 6.2

When regulated expression of a GOI is implemented in transgenic mice, several parameters need to be considered and characterised. For instance, levels of basal expression need to be determined as they might be critical when e.g. a toxic gene is being regulated. Basal expression is influenced by the individual integration site and by the design of the expression cassette. As induction of the autoregulated setting is dependent on a certain basal expression of the system, a detectable level of basal activity is expected (described in chapter 2.2). The nature of the integration site may also restrict expression to certain cell types or tissues which needs to be ascertained. Eventually, times required for complete induction and repression in transgenic animals after administration or removal of dox (kinetics) have to be determined.

Thus, transgenic animals derived from clones 1.1 and 6.2 were administered with dox or left untreated and examined by *in vivo* bioluminescence imaging with the Xenogen IVIS 200 system. Images were analysed concerning the following properties:

a) basal expression in the absence of dox
b) kinetics of dox induced expression
c) reversibility of induction
d) potential tissue restriction of expression.
e) homogeneity of expression

Results for both transgenic lines will be presented together.

Transgenic mice derived from clone 6.2

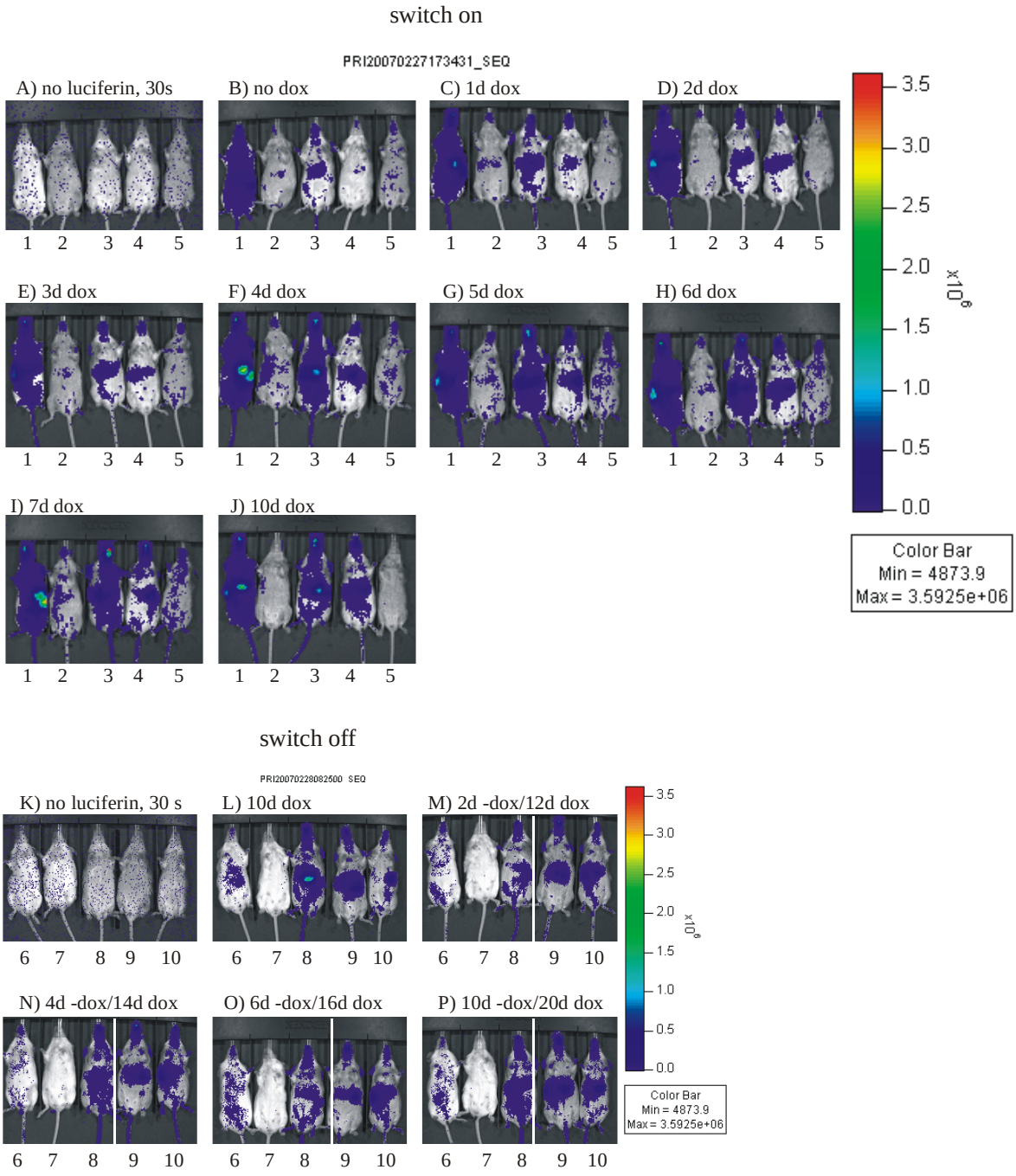


Figure 22: Luciferase expression of transgenic mice derived from clone 6.2
Two groups of 5 transgenic female mice each were kept with/without dox for 10 days (2mg/ml dox in drinking water). *In vivo* BLI was performed on both groups and dox conditions were changed. Mice 1-5 were measured daily for one week after switching on while the switch off group (6-10) was imaged every other day. Mice # 9 and 10 were constantly kept with dox to check luciferase expression over a longer period of time. Coat colour of imaged mice was either light grey or white, depending on the backcross status into Balb/C. Imaging conditions: field of view: 19.6; binning: medium, aperture: f1, exposure time: 2 min if not stated otherwise. Images A) and K) were taken before luciferin injection, showing that no light is emmitted without the substrate. Sensitivity is high in these cases leading to high background activity. Non transgenic control mice are shown in figure 24.

a) A significant level of basal expression could be detected in 2/5 mice in the switch on group, e.g. in mouse 1 (figure 22 panel B). This animal showed a weak signal throughout the whole body while mouse 3 displayed luciferase activity mainly in the stomach region and in the head. Mice 2, 4 and 5 hardly showed any basal luciferase activity at all.

b) When dox was administered via drinking water (switch on), a change could be observed already after 24h (figure 22 panel C): Mouse 1 now emitted a stronger signal from the left upper flank as indicated by a small green focus (cp. rainbow scale). This focus became stronger over the time and seemed to move from one side of the animal to the other while approximately retaining the same height. In panels F, I and J a second signal in this region was discernible. At the same time, also the mouth of animal 1 gave a slightly higher luciferase signal than before. Mice 3 and 4 showed a weak increase in signal intensity after 24h of dox administration. However, the strongest signals were detected after 4d of induction (figure 22 panel F). Mouse 3 exhibited the same pattern as mouse 1 now: mouth and stomach region showed relatively high luciferase activity. Animal 4 did not experience such a strong change, but it did show weak signals from the same parts of the body as did mice 1 and 3 while it had not shown any significant basal expression prior to induction (figure 22 panel B). Mice 2 and 5 could not be induced by dox administration – they exhibited weak, patchy basal expression on some days but appeared to be almost silent the next day (e.g. figure 22 panels C/D; panels I /J).

   The switch off group of mice was kept with dox for 10d and was imaged before and after removal of the inducer. When first imaged, these mice displayed the same pattern as the switch on group after 10d of induction (figure 22 panel L): some animals showed distinct luciferase signals from mouth and upper abdomen (mice 8, 9, 10) while others displayed a patchy, unpredictable expression that was not influenced by dox (mouse 6) or even no expression at all (animal 7). Mice 9 and 10 were kept with dox for 10 more days and retained approximately the same expression levels as before (figure 22 cp. panels L and P). Only mouse 8 could be evaluated for switch off kinetics, as mice 6 and 7 had not expressed significant amounts of luciferase in the first place. Animal 8 lost the expression focus in the stomach region within two days (figure 22 panel M). As expression seemed to be higher again after 4 and 10d of removal of dox (figure 22 panels N and P, respectively), it can be concluded that the

image taken after 2d already showed basal expression levels. Therefore, switch off in clone 6.2 transgenic mice probably is accomplished within 2d.

c) When switched off, mouse 8 reduced luciferase expression levels showing that induction by dox was reversible.

d) To define organs that strongly express luciferase in clone 6.2 derived mice more specifically, animals 1 and 9 were injected with luciferin, sacrificed 3-4min later and their isolated organs were analysed with the Xenogen IVIS system (figure 23). These images demonstrate that inducible luciferase expression is strongest in stomach and brain, which resulted in the signals from upper abdomen and the mouth, respectively, in the whole animal BLI. For animal 1, some small spots of signal could also be detected from the spleen.

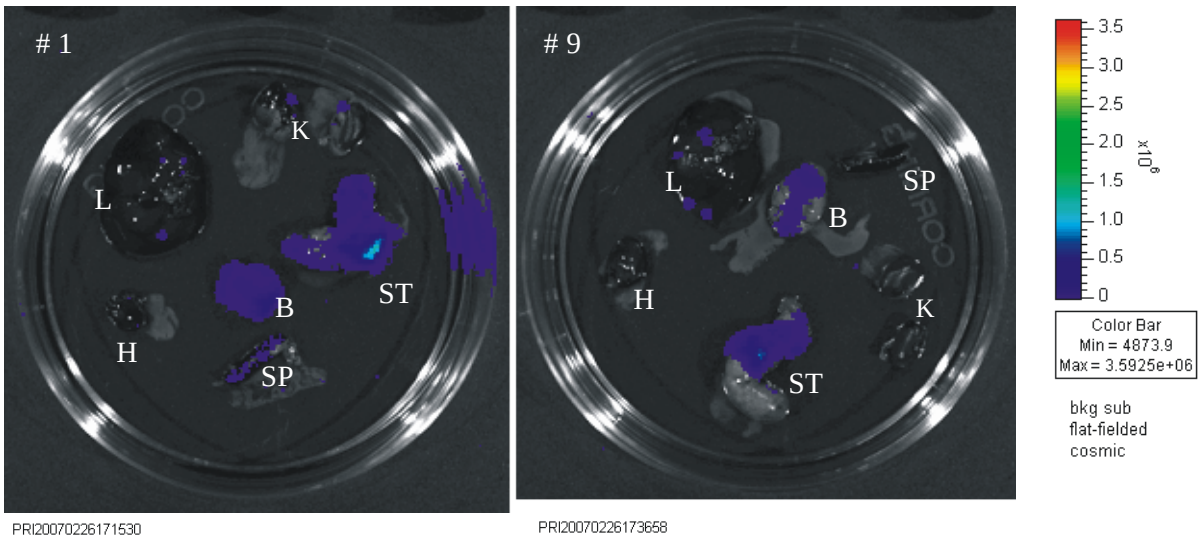


Figure 23: Bioluminescent imaging data form isolated organs of mice 1 and 9
B = brain; H = heart; K = kidneys; L = liver; SP = spleen; ST = stomach; partly opened and cleaned from contents. Imaging conditions: field of view: 13; binning: medium, aperture: f1, exposure time: 2 min

e) Altogether, animal BLI experiments indicate that luciferase expression in clone 6.2 derived mice was heterogeneous – although all animals were transgenic, some of them expressed more luciferase than others (e.g. figure 22 panel I: cp. mice 1 and 5, panel L: cp. mice 8 and 6). Since dox was administered via the drinking water the absolute amount of drug taken up by the animals could not be precisely determined. Calculated to the concentration of dox in the drinking water according to Kistner *et al*. (2mg/ml), mice had imbibed approximately 1 to 1,8 mg dox/day. Thus, it could not be excluded that the heterogeneity of expression was caused by insufficient and/or varying dox uptake. To test whether higher, defined amounts of dox result in higher and more

consistent luciferase expression levels, a group of transgenic mice was administered with 2mg/dose dox twice a day by gavage (figure 24). This mimics plasma levels which are twice as high as those of humans administered with 200mg/d (Zeidner *et al*., 2004).

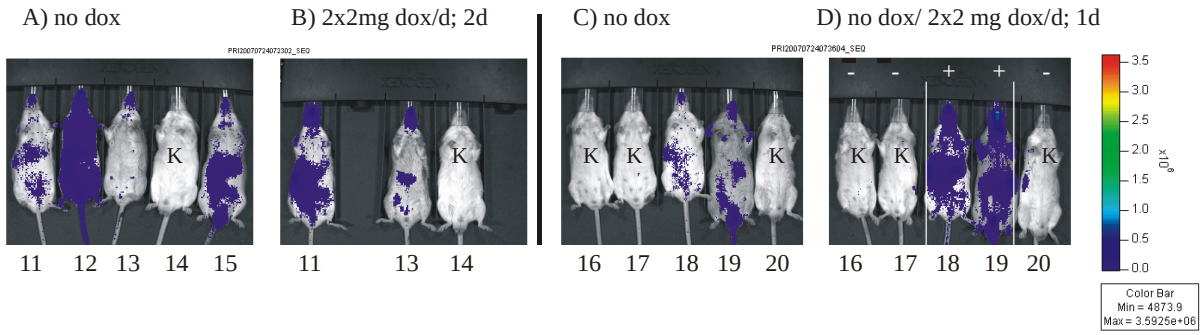


Figure 24: Bioluminescent imaging of clone 6.2 transgenic mice after high dose dox administration (4mg/d)
Mice 12 and 15 died during the experiment due to practical problems. Mice 18 and 19 were therefore fed one day later by gavage. Imaging conditions: field of view: 19.6; binning: medium, aperture: f1, exposure time: 2min. Mice 14, 16, 17 and 20 are non trangenic controls (labelled "K")

Mice 11 and 13 showed clear induction of the autoregulated system after 2d of high dose dox administration (figure 24 cp. panels A and B). However, overall levels were not increased as compared to the first experiments (cp. figure 22 panel D and figure 24 panel B). Even under these controlled conditions still a heterogeneity in expression between individual mice was observed (figure 24 cp. mice 11 and 13 in panel B). Mice 18 and 19 already showed comparatively strong luciferase expression after 24h of dox administration, indicating that expression can be efficiently induced within 1 day in transgenic mice just as in mES cell culture (cp. figure 11).

Transgenic mice derived from clone 1.1

5 clone 1.1 transgenic mice were kept with/without dox each for 10d as described before, analysed once before changing dox conditions and imaged 7 days later (figure 25).

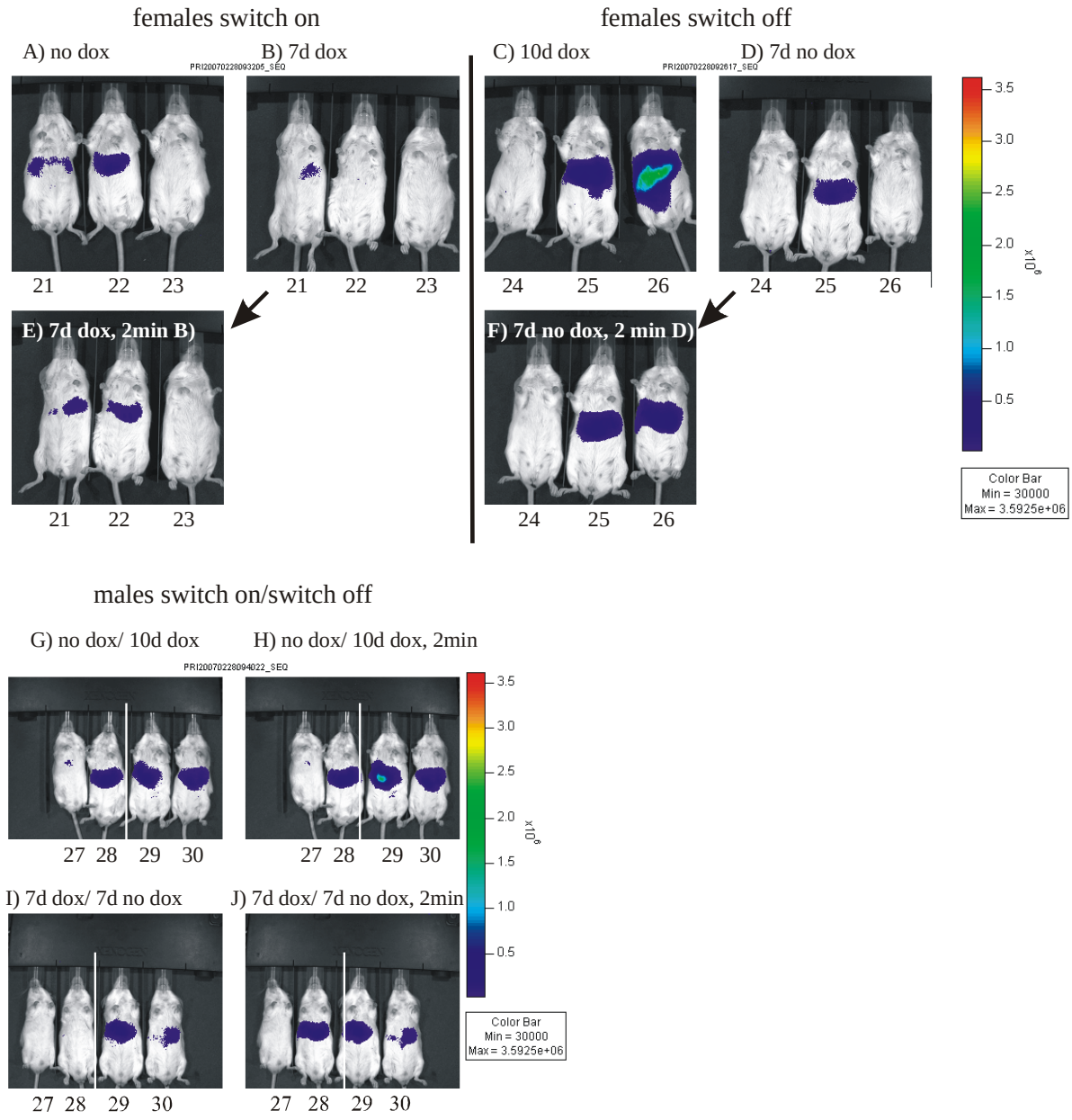


Figure 25: Luciferase expression of transgenic mice derived from clone 1.1
Two groups of 3 transgenic female and 2 male mice each were kept with/without dox for 10d (2mg/ml dox in drinking water). *In vivo* BLI was performed on both groups and dox conditions were changed. Imaging conditions: field of view: 13 for A-F, 19.6 for G-J; binning: medium, aperture: f1, exposure time: 30 s if not stated otherwise. All imaged animals have white coat colour. Non transgenic control mice are shown in figure 24.

a) 3/5 transgenic animals derived from clone 1.1 displayed a detectable level of basal expresssion (figure 25 panel A mice 21 and 22, panel G mouse 28) while in others the cassette was completely silent (panel A: mouse 23 and panel G mouse 27).

b) The switch on group was not influenced by 7d of dox administration (figure 25 cp. mice 21, 22, 23 in panel A with E and mice 27 and 28 in panel G with D). When imaged after induction, the female group even seemed to express less luciferase than before (figure 25 panel B). However, this weak signal might be caused by imaging too soon after luciferin injection, as the acquisition taken a few minutes later demonstrates (figure 25 panel E; the same effect is observed in images D/F, G/H and I/J). Still, exposure time for image E was 4 times longer than for image A, implicating that luciferase expression had not been increased by dox administration. The same is true for the male expressing animal 28 (figure 25 cp. panels H/J).

   The switch off group (mice 24, 25, 26, 29 and 30) in contrast displayed two animals in which luciferase signals are clearly reduced by depletion of dox and another two mice with a slighter yet definite drop of activity (strong decrease: figure 25 mouse 26 panels C/F, mouse 29 panels H/J; weaker decrease: figure 25 mouse 25 panels C/F; mouse 30 panels H/J; note possible differences in exposure times!). The fifth animal of this group (mouse 24) had not shown any luciferase signal at all. Thus, it remained somewhat elusive whether luciferase expression in transgenic mice derived from clone 1.1 is regulated by dox.

c) Consequently, it could not be confirmed whether luciferase expresssion in clone 1.1 transgenic mice is reversible.

d) In all expressing animals, luciferase signals were restricted to the upper abdomen. To determine the organ(s) that gave rise to the reporter signal, mice 22 and 28 were injected with luciferin, sacrificed 3-4m later and their isolated organs imaged immediately. Figure 26 shows that luciferase expression was strictly limited to the liver.

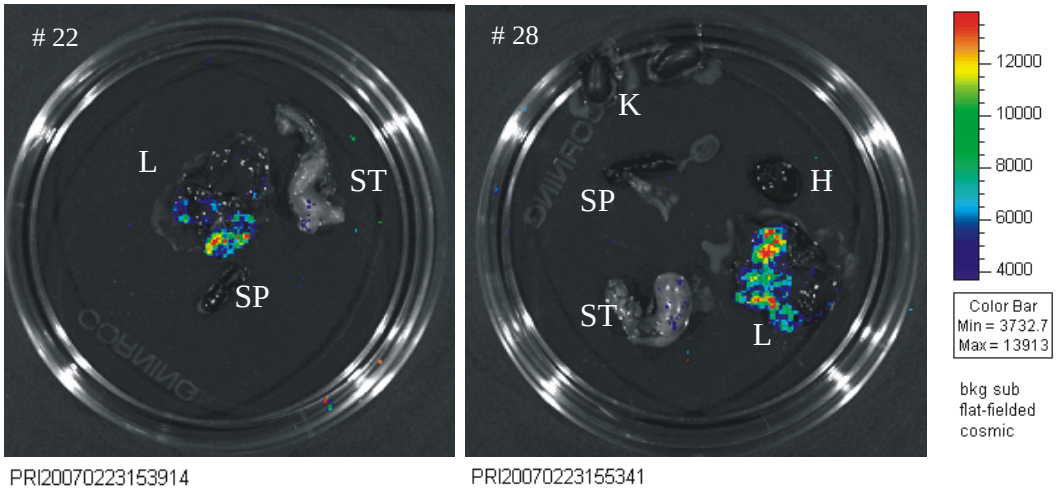


Figure 26: Bioluminescent imaging data form isolated organs of mice 22 and 28
H = heart; K = kidneys; L = liver; SP = spleen; ST = stomach. Imaging conditions: field of view: 13; binning: medium, aperture: f1, exposure time: 2 min. Note that rainbow scale is different and absolute signals are consequently much weaker!

e) Clone 1.1 derived transgenics showed variable expression levels for individual mice: some animals expressed more luciferase than others (e.g. mouse 26 > mouse 25> mouse 24, figure 25 panel C). Thus, luciferase expression in this mouse line was heterogeneous similar to that of clone 6.2 mice.

2.4.1.3 Conclusions: behaviour of the autoregulated cassette in transgenic mice with two different integration sites

Since mice derived from clones 1.1 and 6.2 shared some common features concerning their luciferase expression characteristics, results are discussed together.

1) Basal expression strongly varied between individual animals of the same line: some mice showed hardly any signal whereas others displayed significant luciferase levels in the absence of dox.

2) Expression in clone 6.2 derived mice could be induced to maximum levels within 1-4d by administration of dox. However, this was limited to those mice which had previously shown a high basal activity. The general observation was that the more luciferase activity was detected in the absence of dox, the higher the signal was after induction. Consequently, mice with low or no background did not express more luciferase after subjection to dox. This heterogeneity of expression was not caused by too low amounts of dox in the animals (figure 24).

Transgenics derived of clone 1.1 displayed the same heterogeneity of luciferase expression as described for clone 6.2 animals. However, *in vivo* bioluminescent imaging experiments of transgenic mice and of teratomas did not reliably show regulation by dox.

3) Induction of luciferase expression in clone 6.2 mice was reversible and basal levels were reached within less than 4d after depletion of the inducer.
4) Expression was tissue restricted in both transgenic lines: in clone 1.1 derived mice it was strictly limited to the liver. Although some clone 6.2 animals (e.g. mice 1 and 12, view figure 22 panel J and figure 24 panel A) displayed a signal througout the whole body, high luciferase expression was constrained to the brain and the stomach. In one case, a weak signal could be detected from the spleen (figure 23 organs of mouse 1) and a previous experiment had revealed strong luciferase expression from the kidneys of one animal (data not shown).

2.4.2 In vitro differentiation of clone 6.2 and ROSArtTAautoLuc to hematopoietic progenitors

Directed *in vitro* differentiation of mES cells towards the hematopoietic lineage has been shown to generate progenitors that closely mimic the behaviour of adult hematopoietic stem cells (HSCs) (Pilat *et al*., 2005). The strong resemblance of these so called mES cell derived hematopoietic stem cells (ES-HCs) to HSCs was convincingly demonstrated by their capacity to reestablish the hematopoietic system in lethally irradiated recipient mice.
To extend the scope of possible applications for autoregulated expression of transgenes targeted to precharacterised loci, clones 6.2 and ROSA**rtTA**autoLuc (established by U. Sandhu and S. Bantner) were differentiated to ES-HCs *in vitro* according to Pilat *et al*. (2005). ROSA**rtTA**autoLuc harbours the autoregulated, tet dependent cassette specifically integrated in the ROSA26 locus like the previously described ROSAautoLuc clone. However, an additional reverse transactivator reading frame is inserted under the control of the endogenous ROSA26 promoter (figure 27). This ensures constant levels of rtTA facilitating induction of the positive feedback loop as initiation of expression is independent of basal expression from the autoregulated cassette.

ROSA**rtTA**autoLuc

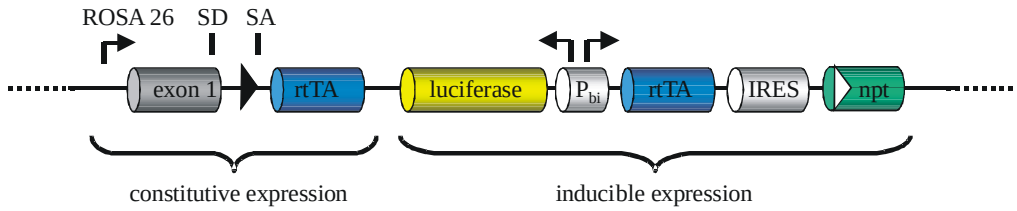


Figure 27: Schematic view of ROSArtTAautoLuc
The previously described autoregulated, tet dependent cassette is integrated in the ROSA26 locus of IB10 mES cells. In addition, the endogenous ROSA26 promoter drives constitutive expression of a second rtTA coding sequence. P_{bi}: bidirectional, tet dependent promoter; rtTA: reverse transactivator of the tet system rtTA2(S)-M2; SD = splice donor; SA = splice acceptor; IRES: Encephalomyocarditis virus internal ribosombal entry site; *npt*: neomycin phosphotransferase gene; FRTwt and FRT5 sites are indicated by black and white arrowheads, repectively.

A detailled description of the applied differentiation protocol is given in “Material and Methods”. Since ectopic overexpression of the human homeodomain transcription factor HoxB4 is required for this differentiation approach, mES cells were first infected with a retroviral vector encoding for HoxB4 and GFP and the pool of expressing cells was recovered by FACS (fluorescence activated cell sorting; described in “Material and Methods”). Infected cells were subjected to the initial differentiation step that involved formation of embryoid bodies (EBs) upon withdrawal of LIF in a methylcellulose matrix. After 6 days EBs were harvested, dissociated and cells were cultured in suspension with medium that contained diverse cytokines needed to gain and sustain ES-HCs (including IL-3, IL-6 and murine Stem Cell Factor).

To evaluate reporter gene expresssion upon induction, dox was added to the media after a certain period of cultivation and cells were harvested and assayed for luciferase activity (figure 28).

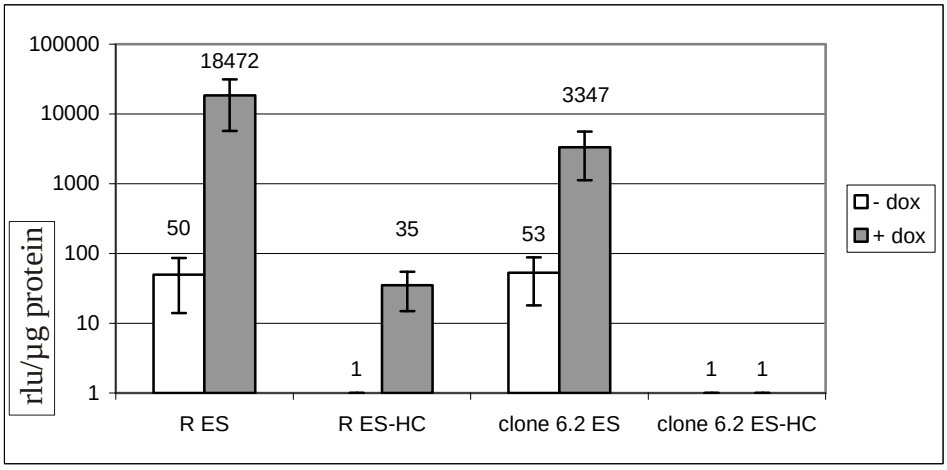


Figure 28: Luciferase activity in ROSArtTAautoLuc and clone 6.2 upon induction on ES cell level and after *in vitro* differentiation to hematopoietic progenitors
R ES: ROSA**rtTA**autoLuc mES cells were kept with/without dox for 2d and harvested for the luciferase assay; sample size for the uninduced/induced states: 4/18; R ES-HC: ROSA**rtTA**autoLuc hematopoietic progenitors were kept in culture for 60d and assayed for luciferase after 2d with/without dox; sample size for the uninduced/induced states: 2/2; clone 6.2 ES: mean values for reporter activity as given in table 3; sample size for the uninduced/induced states: 20/34; clone 6.2 ES-HC: clone 6.2 hematopoietic progenitors were kept in culture for 35d and assayed for luciferase after 2d with/without dox; sample size for the uninduced/induced states: 2/2; IB10 wt control cells did not exhibit any significant reporter levels neither in the ES cell state nor on ES-HC level (data not shown).

Clone 6.2 derived ES-HCs did not not show any detectable luciferase signal in the presence of dox at day 35 of culture just like the IB10 wt control (data not shown). In contrast, ROSA**rtTA**autocLuc even displayed regulated luciferase expression after 60d of culture albeit with a strongly reduced intensity (more than 500 fold less as compared to mean values from ES cells). Interestingly, no basal level of expression could be detected.

To demonstrate the hematopoietic progenitor character of the generated cells, antibody stainings are currently being performed to detect ES-HC/HSC specific markers. In addition, the cells will be transplanted to lethally irradiated, immunodeficient mice. Long term survival of transplanted animals (> 4 months) would prove the cells' potential to completely reconstitute the hematopoietic system, i.e. its "stemcellness".

3 Discussion

The aim of this work was to generate mES cell lines providing high and strictly regulated reporter gene expression. For this purpose, an autoregulated, dox dependent screening cassette was used to randomly tag loci in germline competent mES cells. Concomitantly, heterospecific FRT sites flanking the expression cassette were introduced to enable reuse of the individual locus by subsequent Flpe mediated cassette exchange (RMCE).

Tagged mES cell clones were characterised for regulated reporter gene expression and the number of integrated vector copies. Two clones harbouring a single copy integrate and displaying adequate, regulated reporter levels were surveyed for their expression behaviour after differentiation. To this end, *in vitro* and *in vivo* differentiation assays were carried out which indicated that reporter gene expression of the cassettes in both clones was retained after differentiation. Thus, transgenic mice were established from both lines and further characterised concerning inducibility and expression levels in the whole animal.

Upon improving the selection strategy, targeting of loci in mES cells randomly tagged by an autoregulated cassette was efficiently performed, yielding proof of principle for the overall strategy.

3.1 Flpe mediated cassette exchange of randomly tagged loci in mES cells

Transgene expression in mice is affected by the integration site(s) of the GOI resulting in differing expression patterns in mouse lines derived from individual mES cell clones. Predictable levels of transgene expression may be achieved by integration into defined chromosomal domains. This may be accomplished by application of of BACs, which bring along their own chromosomal environment, or by HR exploiting the known expression properties of endogenous loci. However, the first method only partially solves the problem of reproducible expression (Gong *et al.*, 2003) while the latter suffers from time consuming screening procedures required. Reuse of specific loci in mES cells via RMCE in contrast facilitates generation of genetically modified mice displaying predictable expression patterns and circumventing extensive screening procedures.

Great efforts are currently being undertaken to trap every single gene in mES cells for the generation mouse mutants (IGTC; http://www.genetrap.org). Since the gene trapping consortia implemented RMCE compatible vectors in their trapping strategy, trapped loci may be knocked out or modified to create a range of different alleles at will.

Cobellis *et al*. (2005) showed that trapped genes could be efficiently targeted by RMCE and that reporter expression in mice derived of targeted subclones displayed the same expression pattern as the intact gene.

This “trap and target” strategy is hence well suited to manipulate endogenous genes. However, use of trapped loci to drive transgene expression concomitantly disrupts the endogenous reading frame which might perturb the phenotype that is originally evoked by the transgene.

Thus, the aim of this work was to “tag and target” loci in mES cells which are not required for the host’s gene functions and support regulated transgene expression.

3.1.1 Failure of targeting using the TK/NPT selection strategy

The initial tag and target strategy employed *npt* as a marker for tagged cells and loss of *tk* (i.e. resistance to GCV) for selection for correct targeting events (depicted in figure 17). This negative selection strategy has been realised in NIH3T3 cells by C. Wodarczyk (C. Wodarczyk, 2003), albeit with a relatively high frequency of additional random integration of the targeting cassette. Although experimental factors potentially inhibiting RMCE in clones 1.1 and 6.2 tagged with pTagTK/NPT, like impairment of FRT sites or lack of transfection efficiency could be excluded, no targeted subclones could be isolated. Thus, a low efficiency of this selection strategy in mES cell probably compromised the targeting approach. Consequently, the tagging vector was redesigned to comply with an alternative selection strategy.

3.1.2 The PAC/ΔNPT selection strategy is highly efficient for targeting

The alternative tagging strategy involved screening for tagged loci by puromycin resistance and targeting via complementation of a truncated *npt* gene (figure 19 panel B). 7 puromycin resistant clones generated by electroporation with pTagPAC/ΔNPT with varying levels of luciferase expression were chosen for targeting. 5 of these yielded between 100 and 450 G418 resistant subclones and PCR analysis of 4 subclones each revealed correct targeting events. As the parental clones had not been checked for the number of tagging vectors integrated, it cannot be ruled out that this extremely high efficiency is due to multiple copies available for targeting in each clone. Still, this proves that also the targeting frequency of a single copy will be satisfactory.

Lack of targeting in the remaining 2 clones was possibly caused by single integrations of a 3' deleted vector that was not able to confer G418 resistance upon targeting. In conclusion, this PAC/ΔNPT selection strategy has been demonstrated to be highly efficient and will be the strategy of choice for future tag and target approaches in mES cells.

3.2 Autoregulated, tet dependent reporter gene expression in mES cells, differentiated cells and transgenic mice

The most common methods applied to obtain conditional gene expression in transgenic mice are the genetical switch mediated by the Cre/loxP system and transcriptional regulation driven by the tet system. Since the genetic switch suffers from the prominent disadvantage to be irreversible, tet dependent expression in transgenic mice is gaining more and more attention.

Usually, tet regulation in mice is established by mating an inducible responder line harbouring the GOI driven by a tet dependent promoter to animals expressing the transactivator in a tissue dependent or ubiquitous manner.

A large number of publications describe the succesful application of this "conventional" tet system, i.e. the constitutively expressed transactivator drives expression of the GOI downstream of the tet dependent promoter. These include for instance ablation studies of specific cell types by expression of toxic gene products (Leuchtenberger *et al*., 2001; Lee *et al*., 1998), modeling of pathogenesis during HCV infection on mouse level (Ernst *et al*., 2007), determining the accurate time point when expression of a certain gene is essential for correct embryonic development (Shin *et al*., 1999), or knockdown of endogenous genes by conditional expression of miRNAs (Dickins *et al*., 2007).

While this allows to utilise all transactivator expressing lines established so far to specifically regulate the transgene in various tissues and cell types, two factors make this strategy less attractive: 1) when the transgene is randomly integrated into the genome expression levels of responder lines are unpredictable and can only be evaluated after mating; thus, multiple mouse strains have to be established and mated to ensure generation of a valid model, and 2) breeding is very time-consuming and cost-intensive.

If tissue specificity is not required, these problems may be overcome by an autoregulated design of transgene expression control that allows to generate a complete regulation system in a single step.

Autoregulated expression in transgenic mice subordinating the transactivator to the control of its own dependent promoter has only been utilised by Shockett *et al.* (1995, 2004; configuration as depicted in figure 2 panel A), who did not thoroughly monitor the expression

characteristics. The authors compared luciferase reporter expression data obtained with the conventional Tet_{OFF} system by Furth *et al*. (1994) to autoregulated tTA expression. Furth *et al*. had evaluated inducible reporter gene expression in mice that carried undetermined numbers of CMV promoter driven tTA and tet dependent luciferase gene copies integrated in random loci. Similarly, Shockett *et al*. had randomly integrated undefined copy numbers of tet dependent tTA and luciferase genes into chromosomal DNA. They found that autoregulation yielded approximately two orders of magnitude fold more luciferase activity in thymus and lung as compared to constitutive tTA expression and also detected luciferase in tissues which had previously shown little or no activity with the conventional system (e.g. kidneys and brain). The autoregulated system was used to govern expression of the recombination activating genes 1 and 2 (RAG1 and RAG2) in RAG1 or RAG2 deficient mice, providing inducible activation of V(D)J recombination and the development of primary lymphocytes (Shockett *et al*., 2004).

Results of the studies mentioned above are derived from different mouse lines carrying undefined numbers of the regulatory constructs in different integration sites which strongly curtails comparability. Further, properties of autoregulated expression *in vivo* have not yet been adressed in detail. Thus, it was a major intent of this work to characterise autoregulated tet dependent transgene expression in transgenic mice.

3.2.1 Low efficiency of clone generation by randomly integrating autoregulated, inducible cassettes in mES cells

One of the aims of this work was to identify loci in the genome of mES cells that are supportive to autoregulated expression. To randomly tag loci for subsequent screening, the constructs pTagTK/NPT and pTagPAC/ΔNPT were electroporated into IB10 mES cells. Both vectors share the same autoregulated configuration of tet regulatory components but differ in their selection markers (pTagTK/NPT: $G418^r$, pTagPAC/ΔNPT: $puromycin^r$). A total of 6 electroporations altogether employing ~ 5 x 10^7 cells and 106 µg of pTagTK/NPT only yielded 102 $G418^r$ clones (0,0002 % of total cells; table 2). At the same time parallel control experiments with a vector carrying a constitutively active *npt* gene resulted in 10 fold higher numbers of resistant clones (0.002 % of total cells).

Two possible reasons are conceivable for the low tagging efficiency: possibly tagging was equally efficient as with the constitutive control cassette but transcriptional squelching led to a high percentage of cell death.

Squelching is caused by excessive recruitment of transcription factors for transgene expression and a concomitant overall depletion of these factors for transcription of endogenous genes. Alternatively, demands of (auto)regulated expression on loci in mES cells are higher than for constitutive expression. Performance of the autoregulated cassette presumably is more submitted to influences from chromosomal neighbouring regulatory elements as compared to constitutively expressing constructs. Since mES cells are considered to be “open” for transcription in relation to other cell types, frequent integration of the cassette into transcriptionally silenced genomic sites is not likely a cause for low numbers of resistant clones.

Previous experiments had shown that rtTA2(S)-M2 is capable of reducing transient expression of a cotransfected β-gal reporter plasmid in mES cells in the presence of dox (C. Wodarczyk, 2003). However, expression levels obtained transiently are usually much stronger than expression of integrated constructs. Further, expression levels of pTagTK/NPT are significantly weaker in mES cells as compared to NIH3T3 cells (C. Wodarczyk, 2003). Thus, it is not likely that too high expression of rtTA2(S)-M2 accounts for low clone numbers.

The difference of expression levels between mES cells and NIH3T3 may be explained by the finding that tet dependent promoters contain binding motifs for GATA factors. Gould and Chernajovsky (2004) could show that endogenous GATA factors bind to these motifs suppressing basal promoter activity. Since GATA factors are more abundant in embryonic tissues they may be the cause of suboptimal performance of tet dependent expression in particular in mES cells. In addition, induction of the autoregulated system requires a certain degree of basal expression. If this is abolished by GATA factors, the cassette will be completely silenced.

Analysis of tagged clones revealed that only a minor percentage of cell clones – although being resistant to the selection drug – expressed the reporter gene upon induction with dox. In a certain number of clones this may have been caused by loss of (parts of) the luciferase gene or other essential elements during integration into genomic DNA. Alternatively, low expression levels might have been strong enough to mediate drug resistance whereas luciferase stayed below the detection threshold. This could account for the higher percentage of pTagPAC/ΔNPT tagged, luciferase expressing clones (~20%) in contrast to pTagTK/NPT clones (6%): G418 resistance is already gained when relatively few *npt* molecules are present while *pac* requires a higher expression level concomitantly yielding more clones with detectable luciferase activity.

3.2.2 Reporter levels from the autoregulated cassette strongly fluctuate in mES cells independent of the integration site

Measured luciferase expression levels of mES cell clones carrying the autoregulated cassette showed a standard deviation of at least 66%. This is true for two random integration sites (clones 1.1 and 6.2) as well as for the well characterised ROSA26 locus. In contrast, luciferase constitutively expressed from the endogenous ROSA26 promoter never deviated more than 15%. Consequently, strong fluctuations of expression seem to be an inherent feature of the autoregulated system since they emerge independently of the integration site.

Variations of protein levels between individual, isogenic cells under identical environmental conditions have been observed in both prokaryotes and eukaryotes and are considered to be closely connected to stochasticity of gene expression ("noise"; reviewed by Kaern *et al*., 2005). The manifold causes of cell-to-cell divergence are the subject of ongoing research and can be divided into two classes regarding their sources: a) extrinsic factors such as varying amounts of transcriptional activators and b) intrinsic determinants like the processes of transcription and translation (Raj *et al*., 2006).

To gain detailled insight into the mode of transcription and its implications for gene expression variability, Raj *et al*. (2006) performed *in situ* hybridisation of individual mRNA molecules of both reporter and endogenous genes in CHO cells. The authors found strong variations in the number of mRNA molecules in each cell because the molecules were synthesised in so called "transcriptional bursts". These bursts are dependent on the transition of the respective gene from the inactive to the active state and *vice versa.* Interestingly, two reporter genes expressed from different chromosomal loci displayed no correlation of activation/inactivation whereas both genes integrated at the same site showed simultaneous bursting. These results indicate that fluctuations in mRNA levels are intrinsically random and not significally influenced by global, extrinsic factors. Also, they support the hypothesis that chromatin remodeling is a source of transcriptional bursting. When modeling events during stochastic gene activation and inactivation, Raj *et al*. could show that an increase of transcription factors leads to synthesis of more mRNA molecules per burst rather than enhancing the frequency of bursts.

Fluctuation of chromatin unfolding has been demonstrated to be associated with variable gene expression levels in isogenic cell clones (Sato *et al*., 2004). The authors detected fluctuating expression from a tandem array of 50-60 tet dependent GFP reporter cassettes in presence of constant levels of rtTA and saturating doses of dox. These variations appeared independently of the cell cycle stage and FAC sorted high and low expressing populations returned to a

steady state within less than 3d. Furthermore, FISH (fluorescence *in situ* hybridisation) analyses revealed that levels of GFP expression were correlated to the degree of chromatin unfolding at the integration site of the array.

Regulated gene expression has been shown to follow two distinct patterns, which are not mutually exclusive. In the graded or rheostatic model expression gradually increases upon transcriptional activation in all cells. This applies e.g. when rtTA is constitutively expressed and the level of inducer determines the extent of activation of the transactivator pool (Niwa *et al*., 2000; Biggar and Crabtree, 2001; Bornkamm *et al*., 2005). Bimodal expression in contrast proceeds in a stochastic manner: activated rtTA increases the probability that the regulated transgene will be expressed but does not elevate the expression level as such. This model has been shown to be valid for positive feedback loops (Becksei *et al*., 2001; T. May, 2004). Of note, Isaacs *et al*. (2003) observed that the bimodal expression pattern of a positive feedback can change to a unimodal one depending on the strength of the feedback.

Regarding clones 1.1, 6.2, and ROSAautoLuc, which harbour a strong positive autoregulatory feedback, this means that in the absence of dox the probability of transgene expression is low. However, repression is not completely tight resulting in sporadic activation of transcription and presence low levels of rTA and reporter in the cells. Upon addition of dox, probability of expression is strongly increased which separates the cells into pools of expressing and non expressing cells. When full activation is reached, all cells have shifted to the expressing population. Hence, high luciferase readouts represent a situation when most cells express the cassette whereas low luciferase levels indicate that many cells have stopped transcription. Maybe the cells have an endogenous mechanism that counteracts too high levels of cassette expression caused by the positive feedback loop. Thus, when cells reach a certain peak of expression transcription is stopped and the majority of cells shifts to the inactive state again. After a defined period of recreation however, expression is initiated again. According to this model one would predict an oscillatory mode of expression, where cells increase and decrease frequency of transcription in a synchronised manner. To confirm this, luciferase expression would have to be assayed in short intervals after induction and repression. Oscillations should then be detectable provided half-life times of luciferase mRNA and protein are low enough so that enzyme activity would reflect the actual state of gene expression.

Strong variations of reporter levels might also be related to an unknown factor required for tet dependent expression. If this factor is not homogeneously present it could be limiting for regulated expression and fluctuating reporter levels would consequently mirror its availability.

3.2.3 Hetrogeneity of autoregulated luciferase expression in transgenic mice derived of clones 1.1 and 6.2

Establishing a transgenic mouse line by random integration of a transgene into chromosomal loci strongly suffers from position effects which lead to unpredictable expression patterns. Thus, it is common practice to generate and evaluate several founder lines for each transgenic mouse model. For example, Robertson *et al*. (2002) generated 5 independent mouse lines carrying a tet dependent *lacZ* reporter gene of which only one displayed inducible expression after mating to a tTA expressing mouse strain. This represents the average outcome for transgenic mice created by random integration of the GOI. Since the screening cassette had also been integrated randomly in this work, one would not expect to identify a chromosomal locus that is perfectly suited for regulated transgene expression by analysis of only two reporter lines (i.e. mice derived from clones 1.1 and 6.2).

Luciferase expression levels in mES cell clones 1.1 and 6.2 were well regulated by dox. *In vitro* differentiation experiments and analyses of teratomas induced by these clones indicated that both cell lines retain luciferase expression after diffferentiation and that reporter activity is dependably regulated in the case of clone 6.2 (chapter 2.2.3).

In vivo bioluminescence analysis of transgenic mice derived of clones 1.1 and 6.2 revealed that expression was heterogeneous between individual mice in both lines. This was true for basal expression as well as for the induced state. Unexpectedly, not all mice carrying the autoregulated cassette show basal expression.

Clone 1.1 mice could not be reliably regulated which was documented by the finding that some animals even seemed to express less luciferase after administration of dox. Clone 6.2 mice could be induced by adminsitration of dox provided they had displayed detectable basal levels. This coincides with the notion that only cells with a certain basal level of expression can induce the positive feedback loop. If basal expression is abolished as in the mice without a significant luciferase signal in the absence of dox, expression cannot be triggered but is completely silenced. So, the cassette seems to be shut off in some mice whereas in others it is kept inducible.

Recently, Zhu *et al*. (2007) could show in a classical set-up (constitutive expression of the transactivator) that the bidirectional tet dependent promoter becomes silenced in neurons of mice if it is not active during development. This epigenetic silencing could be avoided by constant basal activity throughout embryonic development or could often be reversed by long-term, high transactivator levels in neurons.

A similar mechanism might be the cause of heterogeneity of expression in the transgenic lines 1.1 and 6.2: stochastic expression of the autoregulated cassette could have led to basal expression in some but not all embryos. Consequently, the tet dependent promoter got silenced in non-expressing embryos and cannot be activated again since rtTA transcription is shut off, too. Thus, only animals with basal expression maintained inducibility.

Silencing of tet driven transgenes is not only a problem in transgenic mice but also requires screening for long term stably expressing clones in cell culture. Pankiewicz *et al*. (2005) investigated the effect of various regulatory proteins targeted to a silenced transgene juxtaposed to a tetO and a Gal4 binding site in cell culture. Interestingly, they found that a Gal4-VP16 fusion was able to reverse silencing whereas tTA (TetR-VP16) had no effect. This suggests that the TetR binding domain cannot interact productively with silent chromatin. When the Gal4 binding domain was fused to mammalian cell activators such as CTF/NF-1 or Sp1, these fusions alone were not capable of inducing silenced expression. However, these factors mediated chromatin remodeling restoring inducibility by tTA. Thus, reversal of epigenetic silencing appears to be a biphasic process that involves transcription factors which remodel chromatin while others subsequently activate transcription. To determine whether heterogeneity of tet dependent expression is due to epigenetic silencing during development in a certain percentage of mice, clone 1.1 and 6.2 mice are currently being bred in presence of dox to ensure constant expression in embryos. Mice from these matings will then be examined by *in vivo* BLI in the presence and absence of dox.

Another factor that might have caused heterogeneous expression is the genetic background of the transgenic mice. Robertson *et al*. (2002) detected a considerable variation of tTA regulated *lacZ* reporter gene expression levels between individual animals that carried the reporter gene in the same chromosomal locus but were of a mixed CBA/Ca x C57BL/6 background. Upon backcrossing into each parental mouse strain expression levels became considerably more homogeneous even after the first backcrossing step. In addition, the authors found that CBA/Ca mice were more permissive to tet regulated transgene expression than C57BL/6J mice. Similarly, clone 1.1 and 6.2 derived mice have a mixed background of Balb/C and 129/Ola strains which could influence heterogeneity of expression. Further backcrossing of 1.1 and 6.2 mice into the Balb/C strain and subsequent analysis of animals will show if expression will become homogeneous in a pure genetic background.

Tissue restriction of expression in both lines (clone 1.1: liver; clone 6.2: brain, stomach) is presumably caused by the individual integration site of the cassette. For example, regulatory

elements or heterochromatin may prevent widespread luciferase expression in all tissues of the animal. Also, the locus of clone 1.1 probably disrupts regulation of transcription.

A recent publication reports that transgenes integrated closely to the telomeres are expressed at low levels in mES cells but are silenced in somatic cells derived thereof (Gao *et al*., 2007). This so called "telomere position effect" involves methylation of DNA and is conceived to be an evidence for epigenetic reprogramming of parts of the genome in the preimplantation embryo. Since the exact integration site of the tagging cassette in clones 1.1 and 6.2 is unknown, impairment of regulated expression in the transgenic animals due to subtelomeric location can not be excluded.

The ROSA26 locus has been frequently used for expression of transgenes driven by either the endogenous or heterologous promoters and is known to permit expression in all tissues unhampered by major disturbances (Zambrowicz *et al*., 1997; Jonnalagadda *et al*., 2005; Strathdee *et al*., 2006; Seibler *et al*., 2007). Thus, it will be highly interesting to analyse animals generated from ROSAautoluc. With these animals, one will be able to clearly distinguish between properties of the autoregulated cassette and influences from the integration site that limit expression in 1.1 and 6.2 animals.

3.2.4 Regulated luciferase expression in hematopoietic progenitors derived by *in vitro* differentiation of mES cells

In vitro differentiation of mES cells provides means to generate genetically manipulated precursors of specific lineages. The major advantages of this strategy are that hitherto no genetic modification as complex as RMCE is feasible in adult stem cells and that the time-consuming generation of a transgenic mouse model may be circumvented.

To test the functionality of the autoregulated cassette and evaluate its potential in *in vitro* differentiation approaches, clone 6.2 and ROSA**rtTA**autoLuc were differentiated to ES-HCs (hematopoietic progenitores derived from mES cells) as described by Pilat *et al*. (2005). Of note, ROSA**rtTA**autoLuc does not only carry the autoregulated cassette integrated in the ROSA26 locus but constitutively expresses a second rtTA reading frame under control of the ROSA26 promoter (figure 27).

Clone 6.2 ES-HCs did not show any detectable luciferase signal after 35d in culture neither in the absence nor in the presence of dox. This may be due to the nature of the integration site of the construct, since *in vivo* bioluminescence analysis of clone 6.2 derived transgenic mice had not revealed distinct expression levels in the hematopoietic system. Luciferase expression specific to the hematopoietic lineage would have yielded strong signals from organs like

spleen and heart which are well supplied with blood (confirmed by *in vivo* BLI analysis of lethally irradiated, surviving mice that had received bone marrow cells from transgenic donors expressing luciferase driven by the ROSA26 promoter, data not shown).

ROSA**rtTA**autoLuc ES-HCs showed weak yet regulated luciferase expression on day 60 in culture. The chromosomal location of the cassette should support expression in blood precursors in this case as Zambrowicz *et al*. (1997) could show that *lacZ* reporter levels are very well detectable in the hematopoietic system if expressed from ROSA26. The more than 500 fold decrease of reporter signal in the induced state as compared to the ES cell state could be connected to the long period of culture. Decrease of transgene expression during culture of mammalian cells is commonly observed when selection pressure is removed, representing a major concern e.g. for the production of recombinant proteins (e.g. Garrick *et al*., 1998; Garrick *et al*., 1996; Migliaccio *et al*., 2000). As discussed previously for the heterogeneous and tissue specific expression pattern in clone 1.1 and 6.2 transgenic mice, epigenetic silencing is a mechanism that is considered to play an important role in the shut down of transgene transcription, especially when the transgene is not expressed (Pankiewicz *et al*., 2005; Zhu *et al*., 2007). However, this effect is dependent on the individual integration site of the transgene and loci have been identified which provide long term, stable expression in the absence of selection pressure (e.g. Schucht *et al*., 2006). Since ROSA26 permits constant expression even in transgenic mice one would assume that this genomic locus is not prone to silencing in cell culture.

Krishnan *et al*. (2006) investigated the impact of epigenetic reporter gene silencing by analysing the effect of a DNA methyltransferase inhibitor (5-azacytidine, 5-aza) and a histone deacetylase inhibitor (trychostatin A, TSA) on CMV promoter driven luciferase expression. For this purpose, the authors transfected the reporter construct into an embryonic rat cardiomyoblast cell line and followed luciferase expression for up to 8 months of culture in the absence of any selection pressure. A progressive decline of luciferase expression could be observed within the highest expressor clone which could partly be rescued by both 5-aza or TSA treatment whereas addition of the transcriptional activator retinoic acid only marginally restored reporter gene expression. The authors further found that after transplantation to rats the cells were twice as long detectable by *in vivo* BLI when they had been cultured in presence of 5-aza beforehand than without any treatment (2 weeks treated vs. 1 week non treated cells). Thus, loss of reporter gene expression due to epigenetic silencing can be a critical factor when tracking stem cell localisation, survival, and differentiation after transplantation in rodents by *in vivo* imaging approaches.

The drop of dox induced luciferase expression during culture of ROSA**rtTA**autoLuc ES-HCs can probably be overcome by constant selection pressure simultaneously forcing expression of luciferase or by treatment with 5-aza or TSA. However, the former alternative precludes regulation of transgene expression while the latter might affect crucial properties of ES-HCs, since the chemicals have a broad impact on cellular characteristics and may lead to undesirable differentiation of cells. Hence, for optimal results, ES-HCs should be used for (transplantation) experiments as early as possible after generation from EBs avoiding long periods of culture.

The fact that inducibility of expression was maintained in ROSA**rtTA**autoLuc ES-HCs although no basal expression was detectable can be explained by the pressence of the additional rtTA reading frame constitutively driven by the ROSA26 promoter. This feature renders regulated expression independent from basal levels to initiate the feedback loop.

3.3 Perspectives

3.3.1 Caveats and improvements

Heterogenic and tissue specific reporter gene expression in mice generated from clones 1.1 and 6.2 is supposed to be caused by the nature of the respective integration sites of the autoregulated cassette. Evaluation of promising loci in mES cell is complicated by the fact that differentiation can have a severe impact on transgene expression characteristics. For example, clone 1.1 had displayed high and well regulated luciferase levels in the ES cell state but lost these properties upon generation of transgenic animals. This once more emphasises the need to identify genomic loci that dependably support regulated transgene expression in somatic tissues.

Thus, a larger number of tagged mES cell clones will have to be screened for an appropriate integration site. This should be done with pTagPAC/ΔNPT since the complementation of the truncated *npt* gene can efficiently be exploited to select for correct targeting events (chapter 2.3.2). Screening procedures will be extended to cells differentiated by teratoma formation because this method successfully mimicked expression in transgenic mice: analysis of teratomas derived from clone 1.1 showed no consistent regulation by dox just like in the clone 1.1 transgenics.

Initiation of the autoregulated positive feed back loop requires basal expression. This has to be kept in mind when designing experiments with this system. While in many cases basal expression will impose no effect, e.g. if transgene levels do not exceed a certain threshold

required for eliciting a reaction, certain biological questions demand strict suppression of the regulated gene. The latter applies for instance when antigens need to be repressed during development of mice to avoid immune tolerance. Mouse models for pathogenesis of HCV and HBV infections are only valid if the mice are not tolerant to the viral proteins due to low level expression during ontogenesis. Low levels of tet regulated viral proteins in the repressed state could not be detected biochemically but still induced immune tolerance in transgenic mice (reviewed by Sun *et al*., 2007). Accordingly, the autoregulated system can be employed for the generation of transgenic mice if the general framework of the experiment agrees with basal exression.

The inherent basal expression of the autoregulated cassette may even have a beneficial effect since it has been shown that activity of a tet regulated transgene can prevent silencing in developing neurons of transgenic mice (Zhu *et al*., 2007).

An issue that might be specific for autoregulated systems is fluctuation of reporter gene expression observed in clones 1.1, 6.2, and ROSAautoLuc. This may be neglected if the response to a regulated transgene is elicited by reaching a certain threshold level and if the lowest induced protein level is sufficient to induce this response. In case the deviating luciferase levels reflect periodic oscillation instead of random fluctuations, this could add another aspect to transgene regulation.

3.3.2 Advantages and possible applications

The autoregulated mode of expression has been shown to principally work in transgenic mice, as luciferase expression in clone 6.2 derived animals that displayed basal activity was inducible and reversible within less than 4d. Limitations like heterogeneous expression levels between mice and loss of regulation (clone 1.1 mice) were most likely due to adverse effects from the cassette's integration site, but were not caused by the autoregulated expression system as such.

Autoregulated expression has several distinct advantages. First, all necessary regulatory elements are combined in a single cassette requiring only a single supportive integration site. Since tagged loci had been screened for their potential to facilitate regulated expression, any cassette – be it constitutive or regulated by another system than tet – should be well expressed after integration into this site by RMCE. Thus, maximal flexibility for the design of expression cassettes is provided.

Another benefit of autoregulated expression is the fact that only very low levels of the transactivator are present in the uninduced state. Long term, high amounts of tet dependent

transactivator have been shown to have adverse effects in certain cells (e.g. Strathdee *et al*., 1999) which has to be avoided in particular in transgenic mice.

The positive feedback loop generated by autoregulation may be exploited to determine activity of moderately active promoters. Instead of directly subjecting a reporter gene to the control of the promoter of interest, a transactivator may be utilised. In addition, a silent autoregulated reporter cassette carrying another copy of the transactivator has to be integrated. If the promoter is active, the first transactivator will trigger initiation of the autoregulated positive feedback loop which in turn will amplify the signal. This principle was put into practice in the design of ROSA**rtTA**autoLuc (depicted in figure 27; the ROSA26 promoter driving the single rtTA gene may be substituted by the promoter of interest).

Targeted integration of transgenes into defined chromosomal sites by RMCE would be of great benefit for the manipulation of adult stem cells as it facilitates predictable gene expression patterns. However, this approach is precluded by the limited time that adult stem cells may be kept in culture and the fact that these cells cannot be cloned. This difficulty can be circumvented by the use of mES cells which may be genetically manipulated by RMCE and subsequently differentiated *in vitro*. By this strategy, not only fully committed somatic cells can be generated but also multipotent progentiors which retain the potential to differentiate into a specific subset of cell types, e.g. hepatic progenitors (Heo *et al*., 2006) or hematopoietic stem cells derived from mES cells (Pilat *et al*., 2005). *In vitro* differentiation to ES cell derived progenitors and subsequent transplantation into recipient mice allows to study the effect of certain transgenes on the respective organ system. Hence, the directed differentiation of modified mES cells circumvents the need to establish transgenic mouse lines and significantly speeds up investigation of gene functions.

The autoregulated cassette is expressed in a stochastic, bimodal manner generating pools of expressing and non expressing cells at intermediate induction. In nature, one would at first expect that such an expression behaviour might have adverse effects on the cells and could even play a role in development of diseases (Kaern *et al*., 2005). Indeed, evidence was found that for expression of essential genes strategies have evolved to reduce the impact of fluctuations indicating the presence of a certain selection pressure on this feature (Fraser *et al*., 2004). However, stochastic gene expression provides a mechanism to generate phenotypically distinct subpopulations from an identical genetic background. For prokaryotes, this mechanism might be exploited to react to a changing environment or sudden stresses. In eukaryotic, multi-cellular organisms stochastic gene expression has been proposed to play a role in differentiation of cells into certain lineages (Fiering *et al*., 2000). If a gene that

determines differentiation into a defined lineage is expressed stochastically, probability of transcription appoints the percentage of cells that will commit to this lineage while all non expressing cells will remain undifferentiated.

So, since fate decisions are of a stochastic nature, an autoregulated transgene expression cassette would perfectly meet the requirements for modeling processes in cell differentiation and development (exemplified in figure 29 by the *in vitro* differentiation to hematopoietic precursors). These processes could then be described by systems biology, extending means to elucidate e.g. signal cascades.

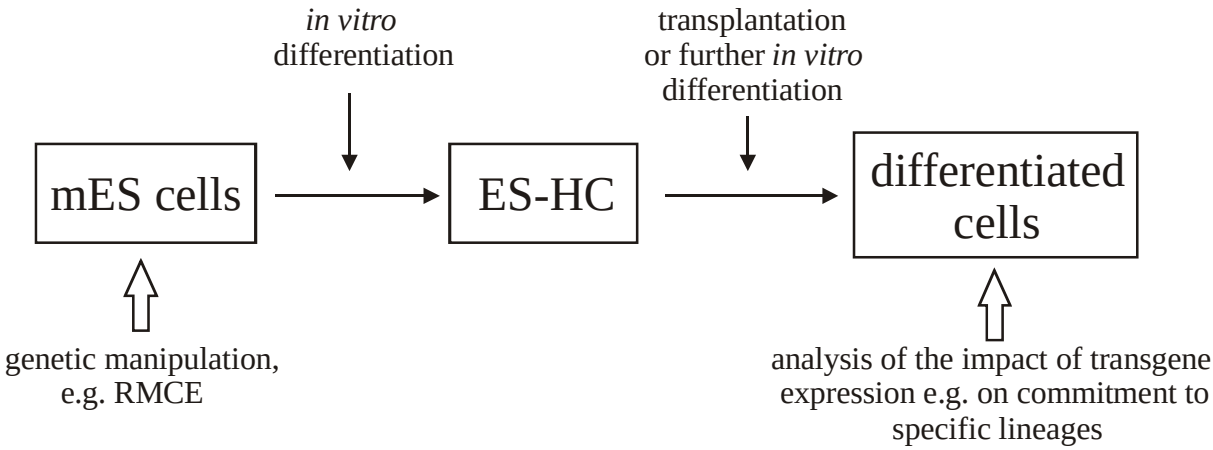


Figure 29: Employment of *in vitro* differentiation and autoregulated gene expression to investigate processes during hematopoietic development
mES cells are first genetically modified, e.g. by introduction of an autoregulated expression cassette into a predefined locus via RMCE. Cells are then differentiated to mES cell derived hematopoietic stem cells (ES-HCs) *in vitro*. Further differentiation may take place either in recipient mice after transplantation or *in vitro* while transgene expression probability may be controlled by the addition of intermediate amounts of inducer. This strategy circumvents the time consuming and expensive generation transgenic mice.

In conclusion, autoregulated tet dependent expression from defined, RMCE targetable loci in mES cells will allow the rapid and simple generation of transgenic mice with conditional transgene expression. In addition, this system permits fast functional studies by *in vitro* differentiation of modified mES cells.

4 Material and Methods

4.1 Equipment

Table top centrifuges	Eppendorf 5417C Heraeus Biofuge 13 Heraeus Christ Minifuge GL Hettich Rontana/S
Cooling centrifuges	Sorvall Superspeed RC5 Minifuge Heraeus-Christ Biofuge fresco inflexible rotors: GSA, GS3, SS34 swing rotor: HB4
Photometer	Nanodrop Spectrophotometer ND-1000, Peqlab
ELISA Reader	Multiskan EX reader, Thermo Electron Corporation
Gelelectrophoresis systems	BRL Horizon 58 BRL Horizon 1114 BRL Horizon 2025 Owl Separation System A2
Power supplies	Desaga Mains Power Supply Unit 1200/200 Biorad Power Pac 300
Microscopes	Leica Labovert FS Nikon TMS
UV-chamber	Hanau
Phosphoimager	Molecular Dynamics Storm 860
Videograph	Biotec Fischer Video densitometer Mitsubishi thermo printer and personal computer
FACS machines & Software	FACSCalibur, Becton Dickinson FACSVantage SE, Becton Dickinson Macintosh Quadra 650
Luminometer	Berthold Lumat LB 9501
In vivo bioluminescence imageing machine	Xenogen IVIS 200 system, Caliper
Cell counter	Casy-DT 1, Schaerfe Systems
Cell culture incubators	Forma Scientifc Water-jacketed Incubator 3336

sterile work benches	Mecaplex, Sterilcard Hood VBM600 and SG400 Heraeus HLB 2448 Heraeus HSP 18
pH Meter	Beckmann M340
Thermomixer	Eppendorf Thermomixer 5436
Vortexer	Scientific Industries Vortex Genie 2
Shaker	Heidolph
Supply of ddH_2O	Millipore MilliQ
Oven for hybridisations	Robbins Scientific Hybridization Incubator 310 Stuart Scientific Hybridization Incubator
Electroporators	Biorad Gene Pulser and Pulse Controller MicroPorator MP-100, Digital Bio
Thermocycler	T3 Thermocycler, Biometra

4.2 Material

Chemicals, enzymes, and cytokines were purchased from the following companies: Abbot, Amersham, Bayer, Biolabs, Biowest, Difco, Flow, Fluka, Gibco, Hoechst, ICN Biomedicals, Invitek, Invitrogene, JRH Biosciences, Macherey-Nagel, Merck, NEB, Peqlab, Pharmacia, Promega, Qiagen, R&D, Roche, Roth, Serva, Sigma, Stratagene, Synchem OHG, and USB. Oligomeres were ordered from MWG and DNA was sequenced by the "Genome Analysis" group headed by Dr. Helmut Blöcker.

Plasticware for culture of eukaryotic cells was ordered from Corning, Costar, Gibco, Greiner, Nunc, Sarstedt, and Seromed.

Texts, graphics and tables of this work were created using the programmes Corel Draw, Word, Power Point, and Excel (Microsoft). Other programmes employed were: Image Quant version 5.0 (Molecular Dynamics) for analysis of autoradiographies; Vector NTI 5.0 (Invitrogen) for the design of primers and cloning strategies; Living image 2.60.1 for analysis of BLI data; and Chromas Version 1.45 (Conor McCarthy, School of Health Science, Griffith University Queensland, Australia) for analysing sequencing results.

4.3 Basic techniques

4.3.1 Sterilisation

Glassware was sterilised by exposition to 180°C for 4h in a cabinet drier. Plasticware like eppendorf cups and pipette tips and solutions were autoclaved for 25min at 121°C.
Temperature sensitive solutions were sterilised by filtering (pore width 0,2µm, Sartorius).

4.3.2 Agarose gelelectrophoresis

1 x TAE-Buffer	40mM Tris/acetate, pH 7,5 ; 20mM NaOAc ; 1mM EDTA
.-	
5 x loading buffer	15% Ficoll, 50mM EDTA, 1 x TAE; 0,05% Bromophenolblue; 0,05% Xylenecyanole

For pouring a 1% gel, 1g of Agarose was dissolved by boiling in 100ml 1 x TAE and 1µl ethidiumbromide (10mg/ml) was added. After the gel had become solid in the tray it was transferred to an electrophoresis chamber filled with 1 x TAE. DNA samples were mixed with 5 x loading buffer and loaded to the gel. For determining the size of DNA fragments, a marker was loaded in parallel (Hyperladder I, Bioline). Electrophoresis was performed at 80-100V and 30mA. Gels were monitored under UV-light (360nm).

4.3.3 Purification of DNA

Purification from Agarose gels:
For preparation of DNA fragments, bands were excised from the gel and DNA was purified with the "QIAquick" gel extraction kit (Qiagen) according to the manufacturer's instructions.

Purification of DNA by Phenole treatment:

Phenole (Roth)	re-distilled, in TE buffer equilibrated phenole (pH 7,8-8,0)
TE buffer	0,1mM EDTA;10mM Tris/HCL, pH 8,0

For removal of proteins the DNA solution was vortexed with the same volume of phenole. The sample was shortly centrifuged to separate the phases and the upper, aeqous phase was transferred to a fresh eppi. The same volume of Chloroforme/Isoamylalcohol (24:1) was added, vortexed and centrifuged again. After transferring the upper phase to a new eppi, this step was repeated to remove residual phenole. The upper phase from the last rinsing step was further treated as described in 4.3.4 to concentrate DNA.

Purification of PCR samples
To purify DNA from PCR samples the "QIAquick" PCR purification kit (Qiagen) was used according to the manufacturer's instructions.

4.3.4 Precipitation of DNA

LiCl/Ethanol	0,6M LiCL in ethanol (stored at –20°C)

Twice the volume of the sample of LiCl/ethanol was added to the DNA solution, shortly mixed and incubated on ice for 10min. The sample was centrifuged (15000 rpm, 15 min, 4°C) and the supernatant was removed. The pellet was washed twice with 70% cold ethanol (with intermediate centrifugation steps) to remove salts. The DNA was finally dried and redissolved as required.

4.4 Modification of DNA

4.4.1 Restriction of DNA

DNA was digested with restriction endonucleases under conditions (buffer, temperature) reccommended by the manufacturer (NEB, Roche). Reactions were stopped by heat inactivation (20min at 65°C/80°C).

4.4.2 Fill in reaction of 5' overhangs

10 x Klenow buffer	50mM Tris/HCL in ddH_2O, pH 7,2; 10mM $MgSO_4$; 0,1mM DTT
dNTP-Mix	1mM of dATP, dCTP, dGTP, and dTTP each

1µg of DNA with overhanging 5' ends was treated with 1 U Klenow enzyme (NEB) in presence of 3µl dNTP-Mix in a total volume of 50µl. The sample was incubated for 30min at 25°C and the reaction was stopped by heat inactivation (20min at 80°C).

4.4.3 Dephosphorylation of DNA fragments

10 x Phosphatase buffer	500mM Tris/HCl, pH 9,0; 10mM $MgCl_2$, 1mM $ZnCl_2$,; 1mM Spermidine

To prevent religation of a restricted vector, 5' and 3' overhanging ends were dephosphorylated by alkaline phosphatase (shrimp alkaline phosphatase, SAP). 20-100pmol DNA with overhanging ends were incubated with 1 x phosphatase buffer and 1-2 U SAP (total volume 50µl) for 30min at 37°C. The reaction was stopped by heat inactivation (20min at 80°C).

4.4.4 Ligation of DNA fragments

5 x Ligase buffer	250mM Tris/HCl, pH7,6; 50mM $MgCl_2$; 25% (w/v) PEG8000; 5mM ATP; 5mM DTT

2µl 5 x ligase buffer, 1 U T4-DNA Ligase (NEB), ~ 20fmol of vector backbone, and ~60fmol of insert were incubated in a total volume of 10µl for > 4h at room temperature (or over night at 16°C). The reaction was stopped by addition of 1µl 0,5M EDTA. The ligation was used for transformation of *E. coli*.

4.4.5 Polymerase Chain Reaction (PCR)

PCR was used to amplify desired DNA sequences. Forward and reverse oligo-nucleotide primers for DNA synthesis were ordered from MWG. PCR comprises a sequence of 3 basic steps that are reapeted 25-35 times ("cycles"): during Denaturation, dsDNA is separated to single strands by high temperatures (94-96°C); Annealing requires lower temperatures for binding of the primers to their complementary sequences; Elongation involves synthesis of complementary strands by the polymerase starting from bound primers.
For standard PCR in this work, the BIOTAQ DNA polymerase from Bioline was used with the provided $MgCl_2$ and NH_4 buffer stocks.

Average PCR sample:

10 x NH_4 buffer	5µl
50mM $MgCl_2$	3µl
10mM forward primer	2,5µl
10mM reverse primer	2,5µl
10mM dNTP Mix (10mM of each dNTP)	4µl
DMSO	1µl
BIOTAQ (2,5 U)	0,5 µl
DNA template	as required
ddH_2O	ad 50µl

Average PCR programme (using T3 Thermocycler, Biometra):

1)	95°C	5min	initial denaturation
2)	95°C	45s	denaturation
3)	~55°C	1min	annealing
4)*	70°C	2min	elongation
5)	70°C	5min	final elongation
6)	4°C	pause	

* steps 2-4 were repeated 30 times before proceeding to step 5.
Annealing temperatures depend on the length and the GC-concentration of primers. Elongation time is dependent on the length of the DNA fragment that needs to be amplified (30s/kb).

4.5 Southern Blot analysis

4.5.1 Generation of radioactively labelled DNA probes using "Rediprime" Kit (Amersham)

Labelling-Mix dATP, dGTP, dTTP; exonuclease free Klenow enzyme; random oligonucleotide primers (9mers; concentrations as provided by the manufacturer)

2,5-100ng DNA (probe for Southern Blot analysis) were diluted in 45µl ddH_2O and denatured in a 95°C hot water bath for 5min. The sample was cooled for 5min on ice and pipetted to the Labelling-Mix (no mixing!). 5µl of α-^{32}P-dCTP (Amersham) were added and the sample was incubated for 30-45min at 37°C after thoroughly mixing. The reaction was stopped by adding 2µl 0,5M EDTA. Labelled DNA was purified from non integrated nucleotides via Micro-Spin G50 columns (Pharmacia Biotech). Loaded columns were centrifuged (2800rpm, 2min) and the probe was again denatured as described above. After cooling on ice it was used for hybridisation. *Npt* and luciferase probes were generated by digestion of pTagTK/NPT as described in the figure legend of the Southern Blot.

4.5.2 Generation of radioactively labelled λ-DNA Marker

10 x dNTP-Mix 25mM of dCTP, dGTP, and dTTP each

To determine the length of the fragments detected by Southern blotting, *EcoRI/HindIII* digested and ^{35}S labelled λ-DNA was used. For labelling, 1µg of digested λ-DNA was mixed with 2µl 10 x dNTP-Mix, 4µl α-^{35}S-dATP (>1000 Ci/mmol) and 2 U Klenow enzyme in a total volume of 20µl. The sample was incubated at 37° for 1h or longer and the reaction was stopped by 20µl 0,5M EDTA. Labelled DNA was purified from non integrated nucleotides via Micro-Spin G50 columns (Pharmacia Biotech). The marker was stored at –20°C.

4.5.3 Isolation of High Molecular Weight (HMW) DNA from mammalian cells/tissues

Isolation of genomic DNA from cells grown on 6-wells:

Modified Bradley's solution 10mM Tris/HCl, pH 7,5 ; 2mM EDTA ; 10mM NaCl ; 0,5% SDS ; 1mg/ml Proteinase K

NaAc/ethanol 75mM Sodium acetate in ethanol

Cells from a confluent 6-well culture dish were washed with PBS and 0,5ml Modified Bradley's were added (proteinase K always freshly added). The sample was transferred to an eppi and incubated at 55°C over night. The next day, 1ml of cold NaAc/ethanol was added and DNA was precipitated for 2-3h at room temperature. The sample was inverted a few times and centrifuged to pellet the DNA (5000rpm, 5min). The supernatant was removed and 0,5ml cold 70% were added and incubated for 30min at room temperature. After centrifugation, this washing step was repeated. The DNA pellet was dried and dissolved in 30-50µl ddH_2O or TE buffer (0,1mM EDTA, 10mM Tris/HCL, pH 8,0).

Isolation of genomic DNA from cells grown on microtiter plates:

Lysis buffer 10mM TRIS/HCl, pH 7,5; 10mM EDTA, 10mM NaCl; 0,5% SDS; 1mg/ml Proteinase K

NaCl/ethanol 150µl 5M NaCl in 10ml ethanol

Confluent microtiter plates were washed twice with PBS and 50µl of the lysis buffer were addded to each well (proteinase K freshly added). The plates were covered, sealed with tape and incubated in a humidified box at 55°C over night. The next day, 100µl cold NaCl were added and the plates were left at room temperature for 1h. They were centrifuged (1000rpm, 5min, 4°C), the supernatant was removed and each well was washed with 150µl cold 70% ethanol. This washing step was repeated (after centrifugation) and the plates were dried. DNA was redissolved in 30µl of ddH_2O or TE buffer.

.

Isolation of DNA from mouse tail tips:

For isolation of genomic DNA from tail tips the DNeasy 96 Blood & Tissue Kit, Qiagen) was used according to the manufacturer's instructions. The DNA dissolved in 400µl ddH_2O was precipitated at 4°C over night by addition of 50µl 4M LiCl and 400µl Isopropanol. Samples were centrifuged (13000rpm, 5min, 4°C) and the pellet was washed twice with 0,5ml 70% ethanol with intermediate centrifugation steps. After drying, the pellet was redissolved in 50µl ddH_2O or TE buffer.

4.5.4 Restriction of HMW DNA for Southern blotting

8-10µg of genomic DNA were digested over night in a total volume of 20µl with 2µl 10 x reaction buffer, 0,5µl RNase (10µg/ml), 0,5µl BSA (100x) and 20-40 U restriction enzyme at the required temperature. Samples were mixed with 5 x loading buffer and loaded to a 0,8% agarose gel. A marker prepared as described in 4.5.2 was also loaded to determine length of fragments after Southern Blotting. Electroporation was performed with 80-100V.

4.5.5 Transfer of DNA to membrane

Electrophoretically separated DNA fragments were transferred from the agarose gel to a positively charged nylon membrane (Hybond-XL, Amersham). The agarose gel was treated with 0,4N NaOH for 10min and put upside down on a tray covered with Whatman paper (3mm). The paper was soaked with 0,4N NaOH and its ends were immersed in a basin with the same solution. The surface of the gel was covered with the wet nylon membrane and two layers of wet Whatman paper were added. A thick layer of absorbent paper tissue and a light weight were put on top. After >6h of DNA transfer from gel to membrane the membrane was neutralised in 2 x SSC and fixed for 2h at 80°C.

4.5.6 Hybridisation of transferred DNA

Hybridisation solution	1M NaCl; 50mM tris/HCl, pH 7,5 ; 10% Dextransulfate ; 1% SDS ; 250µg/ml salmon sperm DNA (sonicated) (stored at –20°C)
20 x SSC	3M NaCl; 0,3M Na_3citrate, pH set to 7,0 with HCl
Wash solution	2 x SSC; 0,5% SDS

20ml of hybridisation solution were thawed, boiled for 5min in a water bath and cooled to 65°C. It was then transferred to a prewarmed hybridisation tube and the membrane was put into the same tube with the DNA carrying side pointing inwards. The radioactively labelled, denatured probe was added and the membrane was hybridised at 65°C over night in a rotating rack. The next day, the membrane was washed 5min with the wash solution at RT and another 30min at 65° while gently shaking. It was kept wet and wrapped in cling film before exposition in an appropriate cassette (Molecular Dynamics). Phospho screens were imaged after > 6h exposition with a Molecular Dynamics Storm 860 Phosphoimager.

4.5.7 Stripping of membranes and rehybridisation

To remove the old probe the membrane was washed with boiling 0,1% SSC/0,5%SDS solution for ~30min. The membrane may be checked for residual probe by exposition and rehybridised as described above.

4.6 Culture and manipulation of *E. coli*

4.6.1 Used strains

XL1-Blue	recA1, endA1, gyrA96, thi-1, hsdR17, suppE44, relA1, lac [F' proAB, $lacI^qZ\Delta M15$, Tn10, (tet^r)]
DH10B	mcrA, mcrB, mmr, hsdR17, deoR, recA1, endA1, lacZDM15

4.6.2 Media

LB-medium	10g/l Bacto-Trypton; 10g/l Bacto-Yeast extract, 5g/l NaCl
Ampicillin	50mg/ml ampicillin in ethanol, filter sterile
SOG-medium	20g/l Bacto-Trypton; 5g/l Bacto-Yeast extract, 10mM NaCl; 2,5mM KCl, 10mM $MgCl_2$; 10mM $MgSO_4$, 20mM glucose ($MgSO_4$ and glucoser were added directly before use.)

Preparation of Agar plates:
15g agar were added to 1l LB-medium and boiled. After cooling the medium to ~45°C, 1ml of the ampicillin stock was added if required and ~30ml were poured into each petri dish.

4.6.3 Electrotransformation of competent bacteria

Generation of electrocompetent bacteria:
E. coli was cultured over night in 10ml LB-medium at 37°C while shaking (180rpm). This pre-culture was added to 1l of LB-medium the next morning and cultivated as described before. When the culture reached an optical density of 0,6 to 0,8 at 600nm (indicating exponential growth), it was centrifuged (3000rpm, 10min, 4°C, GS3-rotor) and resuspended in cold, sterile water. The bacteria were pelleted again by centrifugation and resuspended in 20ml cold, 10% glycerol solution. Another centrifugation step followed (3500rpm, 15min, 4°C, SS34-rotor). The pellet was resuspended in 2-3ml cold, 10% glycerol solution and 50µl aliquots were shock frozen in liquid nitrogen and stored at -70°C. Competence of bacteria may be determined by transformation with a standard plasmid.

Electrotransformation of competent bacteria:
An aliquot of competent bacteria was thawed on ice and ~1µl of DNA or ligation sample were added and mixed. The suspension was transferred to a cooled electroporation cuvette (2mm, Peqlab). Electroporation was performed using a Biorad Gene-pulser with pulse controller at 2,5kV, 25µF and 200Ω. The pulse should last for 4-5ms. The sample was immediately transferred to 1ml SOG-medium and incubated for 30min at 37°C while shaking (180rpm). Aliquots of the suspension were spread on agar plates containing the required selection drug and incubated at 37°C over night for growth of resistant colonies.

4.6.4 Preservation of *E. coli* strains

For short term storage *E.coli* was expanded over night at 37°C on agar plates and subsequently kept at 4°C. For long-term storage 500µl of suspension culture were mixed with 500µl 87% glycerol solution and stored in glass tubes at -70°C.

4.6.5 Small scale plasmid DNA isolation for analysis

STET buffer	80g/l Sucrose; 0,5% Triton X100; 50mM EDTA; 10mM tris/HCl pH8,0
TE buffer	0,1mM EDTA;10mM Tris/HCL, pH 8,0
Lysozyme	10mg/ml Lysozyme in TE buffer
Ammonium acetate	8M NH_4OAc
TE RNase	10µg/ml RNase A in TE buffer

2ml of LB-medium containing 50µg/ml ampicillin were inoculated with the respective *E.coli* clone and cultured over night at standard conditions. Suspensions were transferred to eppis, centrifuged (13000rpm, 1min) and the supernatant was discarded. Pellets were resuspended in 500µl STET buffer. After addition of 50µl lysozyme solution samples were incubated for 2-3min at room temperature. The reaction was stopped by heat inactivation (95°C, 90s) and the eppis were centrifuged (13000rpm, 1min). The viscous pellets were removed with toothpicks and 50µl ammonium acetate and 500µl Isopropanol were added to the lysate. Samples were again centrifuged (13000rpm, 1min) to precipitate the DNA, the supernatant was removed and pellets were dried. The DNA was redissolved in 40-50µl of TE RNase.

4.6.6 Large scale plasmid DNA isolation (Qiagen Plasmid Maxi Kit)

For isolation of large amounts of plasmid DNA, the Plasmid Maxi Kit from Qiagen was used with the provided solutions according to the manufacturer's instructions. For this purpose, 200ml LB-medium containing the appropriate selection drug were inoculated with the desired *E.coli* clone and incubated over night (37°C, 180rpm). The suspension was pelleted by centrifugation (3000rpm, 10min, 4°C, GS3-rotor) and treated as described in the kit's manual. DNA was dissolved in adequate amounts of TE buffer.

4.7 Culture and manipulation of eukaryotic cells

4.7.1 Cells and cell lines used

IB10	Subclone of murine embryonic stem cell line E14 derived from 129/Ola mouse strain blastocysts (Robanus-Maandag *et al.*, 1998; Hooper *et al.*, 1987)
MEFs	primary embryonic mouse fibroblasts isolated from different mouse strains (129/Ola x Balb/C chimaeric embryos for analysis; C57BL/6J for feeder cells supporting mES cell culture)
293EBNA-LIF	HEK293T cells stably expressing Leukaemia Inhibitory Factor (LIF) and recombinant LIF were provided by Dr. Martin Hafner (cloning and cell culture) and Dr. Werner Müller from the department of "Experimental Immunology" (design of producer cell line employing EBNA system)
HEK-293T	human embryonic kidney cell line transformed by adenovirus type 5 (DSMZ ACC 110), into which the temperature sensitive gene for SV40 T-antigen was inserted

4.7.2 Media components

DMEM (Dulbecco's Modified Eagle's Medium)	for mES cells: DMEM+GlutaMAX-I (Gibco) for differentiated cells: 13,63 g/l DMEM powder (Sigma, already containing sodium pyruvate, high glucose) 44 mM $NaHCO_3$ 10 mM HEPES pH 7,2
PBS (phosphate buffered saline)	140 mM NaCL 27 mM KCl 7,2 mM Na_2HPO_4 14,7 mM KH_2PO_4 pH 6,8-7,0
TEP (trypsin EDTA)	6 mM EDTA; 0,1 % trypsin (Sigma) in PBS
100 x Pen/Strep	6,06 mg/ml ampicillin (10000 U/ml), 10 mg/ml streptomycin; set pH to 7,4 by addition of NaOH for dissolving (stored at –20°C)
100 x Glutamine	29,23 mg/ml Glutamine (stored at –20°C)
Gelatin solution	2% stock solution Type B: from bovine skin (Sigma); diluted to 0,1% (w/v) with PBS for use (stored at 4°C)
FCS (Fetal Calf Serum)	for mES cells: ICN Biomedicals GmbH for other cell lines: JRH Bioscience, Biowest
LIF (Leukaemia Inhibitory Factor)	500-800µl/500ml supernatant; produced as described in 4.7.5
100 x Sodium Pyruvate	100 mM sodium pyruvate (Gibco)
100 x Nonessential Amino Acids	Gibco
1000 x β-Mercaptoethanol	100 mM β-mercaptoethanol (7,014µl/ml in ddH_2O, filter sterile; stored at –20°C)
Doxycycline	2mg/ml doxycycline-hyclate (Sigma) in 70% ethanol, filter sterile (stored at –20°C)
G418	100mg/ml G418 in ddH_2O, filter sterile (stored at –20°C)
Puromycin	5mg/ml puromycin in ddH_2O, filter sterile (stored at –20°C)
Ganciclovir	10 mM ganciclovir in H_2O (stored at 4°C)

4.7.3 Media composition

for mES cells:	1 x Pen/Strep 1 x Sodium Pyruvate 1 x Nonessential amino acids 1 x β-Mercaptoethanol 15% FCS (heat inactivated: 30 min at 56°C) 500-800µl LIF/500ml DMEM+GlutaMAX-I (Gibco)
for MEFs: ("5+" medium)	1 x Pen/Strep 1 x Glutamine 1 x nonessential amino acids 1 x β-Mercaptoethanol 10% FCS DMEM
for HEK-293 derived cells: ("3+" medium)	1 x Pen/Strep 1 x Glutamine 10% FCS DMEM

4.7.4 Culture of mammalian cells

mES cells were grown on suitable culture dishes either coated with gelatin or with feeder cells (preparation described in 4.7.6.1). For gelatin coating, a 0,1% gelatin solution was added so that it completely covered the surface of the dish. After incubation for > 15min at 37°C excess of gelatin was aspirated and dishes were used for culture. mES cells were kept at 37°C and 7% CO_2 in humidifying incubators. The medium was changed daily and cells were split at ~ 75% confluence (every 2-3d).
All other cells/cell lines were kept on untreated cell culture dishes at 37°C and 5% CO_2 at maximal humidity. Medium was changed every 3-4d and cells were passaged at confluence.
For passageing, cells were washed with PBS and incubated with TEP for < 3min at 37°C. Detached cells were taken up in fresh medium and transferred to new culture dishes.
For determining the number of suspended cells an aliquot of the suspension was counted with a CASY-1 DT cell counter (Scharefe Systems, Germany).

4.7.5 Production of LIF

Culture supernatant of LIF expressing 293EBNA-LIF cells was either produced by fermenter culture by Dr. Mariela Bollati Fogolin (formerly group of "Experimental Immunology" headed by Dr. Werner Müller) or by conventional cell culture. For the latter, cells were grown on big culture flasks in the presence of 0,3µg/ml puromycin. At ~ 80% confluence, the medium was changed removing puromycin pressure. The supernatant was harvested after 3 d, sterile filtered and stored at - 20°C. Concentrations of LIF required for mES cell culture were determined by a growth assay culturing mES cells in presence of different dilutions of the supernatant. mES cells were assayed for appropriate morphology and the second lowest LIF concentration that produced undifferentiated mES colonies was used for preparing mES cell media (500-800µl/500ml).

4.7.6 Preparation of Mouse Embryonic Fibroblasts (MEFs)

For isolating MEFs, pregnant mice were sacrificed at day 13,5 by cervical dislocation and disinfected with 70% ethanol. Embryos were extracted and transferred to a dish with PBS. The amniotic sac, the head and the blood building organs from each embryo were removed using forceps and a scalpel. The embryos were then chopped up to very small pieces which were taken up in 5ml of TEP and transferred to 15ml tubes. During an incubation time of 30min at 37°C, tubes were thoroughly shaken every 5min. The suspension was transferred to 50ml tubes and 30ml of medium were added for washing. Cells were pelleted (1000rpm, 5min) and the supernatant was aspirated. The washing step was repeated and the cells obtained from one embryo were finally seeded to two 10cm culture dishes.

4.7.6.1 Use of MEFs as feeder cells for mES cell culture

For use as feeder cells supporting culture of mES cells, MEFs were expanded on 25 cm dishes for 4-5 passages, harvested by TEP and irradiated with ~24,8Gy (700s at a dose rate of 127,6 Gy/h). Irradiation inhibits further cell division preventing potential overgrowth of feeders when coculturing mES cells. Irradiated feeders were cryoconserved for later use. When needed, they were thawed and seeded at a density of ~ $2{,}5 \times 10^4$ cells/cm^2.

4.7.7 Long term storage of mammalian cells

Cells were harvested with TEP and pelleted at 1000rpm for 5min. The pellet was resuspended in 0,5-1ml freeze medium (mES cells: 90% FCS, 10% DMSO; all other cells: 95% FCS, 5% DMSO) and transferred to cryo vials (Corning). For freezing cells a slow decrease of temperature is required. Thus, mES cells were frozen in isopropanol filled containers at-70°C. After > 1d they were transferred to liquid nitrogen for long term storage. All other cells were kept on ice for 30min and at –70°C for 24h prior to storage in liquid nitrogen. Thawing was performed fast in a 37°C water bath and cells were washed with medium to remove residual DMSO. After centrifugation (100rpm, 5min) and aspiration of the supernatant they were seeded to fresh cell culture dishes.

4.7.8 Production of HoxB4 transmitting viral particles

HoxB4 transmitting retroviral particles pseudotyped with VSV envelope protein were transiently produced by calcium phosphate transfection of the required plasmids into HEK-293T cells as described in 4.7.9.2. 10^7 cells were seeded to a 25cm culture dish and transfected the following day with 2 ml $CaCL_2$/HEBSbuffer containing 60µg of pSVGP1, 7µg of pVSV-G, and 20µg pHoxB4-2A in a total volume of 15 ml medium. Medium was exchanged for 10ml fresh medium containing 20mM HEPES after 6-12h and the supernatant was harvested twice at 48 and 72h post transfection. It was then filtered (pore width 0,45µm) and stored at –20°C.

4.7.9 Gene transfer methods

4.7.9.1 Infection of mES cells with pHoxB4 retroviral particles

For transduction of mES cells with HoxB4 retroviral particles 10^6 cells were seeded to a gelatinised 10 cm dish and infected with a multiplicity of infection (moi) < 0,01 (i.e. less than 0,01 infectious particles/cell) on the following day.
The moi of the respective supernatant batch (produced as described in 4.7.8) was determined by test infections of mES cells with serial dilutions and subsequent FACS analysis (pHoxB4 coexpresses eGFP).
8µg/ml polybrene were mixed with the concentrated viral supernatant which was then added to the cells (5ml on 10 cm dishes; 1ml on 6-wells). To reduce the probability of differentiation of mES cells, adequate amounts of LIF were added. After 10-18h cells were split to feeder

coated dishes and the pool of eGFP expressing mES cells was isolated by FACS 2d post infection (4.7.11).

4.7.9.2 Transfection by calcium phosphate/DNA precipitation

2 x HEBS buffer	280mM NaCl, 50mM HEPES, 1,5mM Na_2HPO_4, pH 7,1 (filter sterile)
$CaCl_2$ solution	2,5 M $CaCl_2$ in ddH_2O (filter sterile)

5 x 10^5 mES cells were seeded to gelatinised 6-wells for targeting approaches by $CaPO_4$ precipitation. Similarly, HEK-293T cells were seeded as described in 4.7.8 for the production of retroviral particles. The next day, the DNA(s) to be transfected were mixed with the $CaCl_2$ solution in a 5ml tube (15µl 2,5M $CaCl_2$, x µl DNA, ddH_2O ad 150µl for one 6-well; scale up to 1ml for a 25cm dish). The same amount (i.e. 150µl and 1ml, respectively) of 2 x HEBS buffer was slowly added while vortexing the tube. The mixture was incubated at room temperature for 3-10min and then added to the culture dish. Medium was changed after 6-12h. If needed, cells were induced with 2µg/ml dox after one day and selection pressure was imposed 2d post transfection.

4.7.9.3 Electroporation of mES cells

Plasmid DNA was digested and purified as indicated in tables 1 and 6. 10^6-10^7 mES cells were harvested per electroporation, centrifuged (1000rpm, 5min) and the pellet was washed with PBS to remove residual medium. For electroporation with the Gene Pulser (Biorad) cells were resuspended in 500-700µl PBS, the DNA was added and the sample was transferred to an electroporation cuvette (4mm, Peqlab). Electroporations were performed under conditions indicated in tables 1 and 6. When applying the Microporator technology (Digital Bio), 100µl electroporation tips were used according to the manufacturer's instructions (details see table 6). After electroporation, cells were taken up in fresh medium containing 2µg/ml dox and seeded to feeder coated 10 cm dishes. Selection was started 2d after electroporation by either 0,4mg/ml G418 or 1µg/ml puromycin in the presence of dox. Resistant clones were picked after 9-14d.

4.7.9.4 Transfection using commercial transfection reagents

5 x 10^5 mES cells were seeded to gelatinised 6-wells. GenePorter^{TM}2 (Peqlab), FuGENE6 (Roche) and Lipofectamine2000 (Invitrogen) were used for DNA transfer according to the manufacturers' instructions. Of note, Lipofectamine2000 yielded the highest transfection efficiency in mES cells (7µl Lipofectamine2000/transfected 6-well). For targeting using Lipofectamine2000, 2,5µg of the Flpe expression vector and 1,5µg of the targeting vector were cotransfected. Medium was removed after 6-12h and cells were passaged to feeder coated 10 cm dishes in the presence of 2µg/ml dox. Selection was started 2d post transfection by either 10µM GCV and 0,4mg/ml G418 (pTagTK/NPT targeting) or solely 0,4mg/ml G418 (pTagPAC/ΔNPT targeting) in presence of dox. Resistant subclones were picked after 9-14d.

4.7.10 Isolation of selected clones

Colonies of drug resistant clones were marked on the bottom of the culture dish. The plate was washed with PBS and a small amount of fresh PBS was added that covered the surface of the dish. The plate was put under a microscope in the clean bench and colonies representing cell clones were gently removed with a 20µl Gilson pipette. Cells were transferred to a microtiter dish, digeseted with 20µl of TEP for 1-2min, taken up in 150µl medium and transferred to fresh microtiter wells from which they were expanded..

4.7.11 Fluorescence Activated Cell Sorting (FACS)

FACS was used to determine titers of eGFP producing viral particles, to sort for pHoxB4-2A infected mES cells and to test co-transfection efficiency of two plasmids (eGFP and dsRed expression vectors). For FACS, cells were washed with PBS, harvested with TEP and resuspended in FACS buffer (PBS with 2% FCS) after centrifugation (1000rpm, 5min). Dead cells were stained by 50µg/ml propidium iodide when required. Analysis of cells was performed with a FACSCalibur machine (Becton Dickinson). Debris was eliminated from analysis by setting the forward scatter < 200 as were dead cells by excluding propidium iodide positive signals. Sorting of eGFP positive cells was done with a FACSVantage SE (Becton Dickinson).

4.7.12 In vitro differentiation of mES cells

4.7.12.1 Unspecific *in vitro* differentiation

Undirected differentiation of mES cells was performed by seeding ~7,5 x 10^5 cells in 13ml of 3+ medium to bacterial dishes to prevent cells from adhering to the culture dish surface. Suspension culture and concomitant withdrawal of LIF led to the formation of embryoid bodies (EBs) which to a certain degree mimic developmental processes of a normal embryo. EBs were cultured for 5d, harvested by centrifugation (500rpm, 5 min) and plated to gelatinised 10cm dishes. There, they attached to the surface of the dish and formed so called "outgrowths" of differentiated cells. Cells were dissociated by TEP and split to gelatinised 6-wells 4d later. They were cultured for another 4d in the presence or absence of dox. Finally, samples were harvested for analysis of luciferase activity.

4.7.12.2 In vitro differentiation to hematopoietic progenitors

IMDM (Iscove's modified Dulbecco's medium)	IMDM powder as supplied by Gibco (already containing glutamine and 20mM HEPES) 30,2g $NaHCO_3$ ddH_2O ad 10l pH 7,2
IMDM based mES cell medium	1 x Pen/Strep 1 x Sodium Pyruvate 1 x Nonessential amino acids 1 x β-Mercaptoethanol 15% FCS (heat inactivated: 30 min at 56°C) 500-800µl LIF/500ml IMDM
2 x Methylcellulose stocks (prepared as described in 4.7.11.3)	20g methylcellulose (Fluka) IMDM powder 3,025g $NaHCO_3$ 1l ddH_2O

Diff-medium	1 x Pen/Strep 1 x Glutamine 15% FCS 5% PFHM-II (protein free hybridoma medium, Gibco) 300µg/ml human holo-transferrin (Sigma) 0,4mM MTG (monothioglycerol) 50µg/ml ascorbic acid 50% 2 x Methylcellulose stock IMDM
SCM (Stem Cell Medium for culture of ES-HCs) (prepared as described in 4.7.11.4)	1 x Pen/Strep 100ng/ml mSCF (murine Stem cell factor,Sigma) 2ng/ml mIL3 (murine interleukin 3, R&D) 10ng/ml Flt3-L (murine Flt3 ligand, R&D) 5ng/ml hIL6 (human interleukin 6, R&D) 40ng/ml Long R^3 IGF-1 (recombinant insulin-like growth factor analogue, Sigma) 1µM dexamethasone (Sigma) StemPro34 + Nutrient supplement (Gibco)

In vitro differentiation of mES cells to ES-HCs (hematopoietic progenitors derived *in vitro* from mES cells) was performed according to Pilat *et al.* (2005). Pools of mES cells overexpressing HoxB4 (established as described in 4.7.8, 4.7.9.1, 4.7.11) were cultured for 2d in IMDM based ES cell medium. Cells were harvested with TEP, centrifuged (1000rpm, 5min) and washed with IMDM ES cell medium without LIF to remove any residual LIF that might disturb differentiation. 10^5 cells were taken up in < 500µl of IMDM-ES (without LIF) and pipetted on top of 20ml warm Diff-medium in a 50ml tube. The tube was vortexed to distribute the cells in the viscous medium. It was kept in a 37°C water bath for a few minutes until the majority of air bubbles had disappeared. The suspension was plated in a spiral on a 25cm bacterial dish and the medium was dispersed by gently tilting the plate. The dishes were put into a box with a lose lid, wrapped in a plastic bag with small holes and areated for a few seconds with CO_2 (this was done whenever handling the cells). The box was then kept for EB formation at 37°C, 7% CO_2 and maximal humidity. On day 3 of differentiation, 10ml of fresh Diff-medium were added and dispersed by gently tilting the dish. Medium was changed on day 5: the plate was tilted and the suspension was carefully harvested with a plastic pipette (wide aperture prevents destruction of EBs) and transferred to a 50ml tube. The dish was washed with warm IMDM to recover all EBs. The tube was centrifuged at low speed (300rpm, 5min) and the supernatant was removed. EBs were carefully resuspended in fresh Diff-medium and replated using plastic pipettes. On day 6 or 7, EBs were harvested as described before and dissociated. For this purpose, the EB pellet was washed with PBS and 3ml of TEP were added. The tube was incubated in a 37°C water bath and dissociation was stopped by addition of 3µ prewarmed FCS as soon as the suspension became cloudy (i.e. disaggregation of EBs). Cells were vigorously pipetted up and down, centrifuged (1000rpm, 5min) and washed with PBS to completely remove FCS. A single cell suspension was obtained by using a cell strainer (70µm, Becton Dickinson). Cells were counted, centrifuged (1000rpm, 5min) and the pellet was resuspended in SCM and transferred to suspension 6-well plates. Cells were kept at a density of at least 10^6 cells/ml and debris and terminally differentiated cells were removed from the culture by Ficoll centrifugation (4.7.12.5) 2-3 times per week. At least 50% of the medium was changed every day. Control cells not expressing HoxB4 stopped growing after approximately 35d.

4.7.12.3 Preparation of 2x methylcellulose stocks

For preparation of 1l 2 x methylcellulose stock one sterile erlenmeyer flask was weighed and filled with 450 ml ddH_2O. The water was boiled on a heater and 20g methylcellulose (Fluka) were added. This was simmered for 10min while being mixed with a magnetic stirrer, boiled up shortly for dissolving and sterilising and let cool down to room temperature. IMDM powder for 1l and 3,025g $NaHCO_3$ were dissolved in 450ml ddH_2O and filtered sterile. This 2x medium was poured to the methylcellulose while stirring. ddH_2O was added to a final weight of 1000g (- weight of flask). The stock was moved to 4°C and stirred over night. The next day, aliquots of ~40ml were poured into 50ml tubes and shock frozen in liquid nitrogen. The aliquots were stored at –70°C. For preparation of Diff-medium an adequate number of stocks was thawed over night at 4°C and centrifuged (2000rpm, 10min). 5ml of each methylcellulose stock were left in the tube to get rid of fibrous cellulose debris.

4.7.12.4 Preparation of SCM (stem cell medium for culture of ES-HCs)

500ml StemPro34 was mixed with the 13ml of Nutrient supplement (Gibco) and adequate amounts of IGF-1 (dissolved as reccomended by the manufacturer), dexamethasone (1mM stock in 96% ethanol) and Pen/Strep were added as given in 4.7.12.2. After mixing, 45ml were aliquoted into 50ml tubes and shock frozen in liquid nitogen. These stocks were stored at –70°C. When needed, stocks were thawed over night at 4°C and cytokines were added freshly. Cytokine stocks had been prepared by dissolving the lyophilised powders in 0,1% BSA in PBS and aliquots were stored at –70°C.

cytokine	stock concentration	amount added to 45ml SCM
mIL3	20µg/ml	4,5µl
hIL6	20µg/ml	11,25µl
Flt3-L	100µg/ml	4,5µl
mSCF	100µg/ml	45µl

4.7.12.5 Ficoll gradient centrifugation

ES-HC cultures were cleared from debris and terminally differentiated cells by Ficoll density centrifugation 2-3 times per week. Cell suspensions were transferred to 15ml tubes. 3-5ml of cold Ficoll (Biochrom) were carefully pipetted under the cell suspension and the tubes were centrifuged (2100rpm, 10min). Upper phase and interphase containing the healthy cells were transferred to fresh tubes and centrifuged again (1200rpm, 5min). The pellet was resuspended in fresh SCM.

4.8 Protein analysis

4.8.1 Extraction of proteins from cells

To determine expression levels of firefly luciferase all water soluble proteins had to be extracted from the cells. For this purpose, cells were harvested from 6-wells and the cell pellet was resuspended in 50µl Tris-HCL (pH 7,6) by vortexing. Cells were broken by 4 repeated freeze-and-thaw cycles in liquid nitrogen and a 37°C water bath, respectively. After centrifugation (15000 rpm, 10min, 4°C) the supernatant was used for luciferase and BCA assays.

4.8.2 Luciferase assay

Firefly luciferase catalyses oxidation of D-luciferin to oxyluciferin. Detection of concomitantly emitted photons allows to determine expression levels of the luciferase reporter gene.

Luciferase buffer	25mM glycylglycine, 15mM $MgSO_4$ in ddH_2O, pH 7,8 (stored at 4°C)
ATP solution	5mM ATP in ddH_2O, pH7,5 (stored at –20°C)
Luciferin solution	0,1mM synthetic D-luciferin (Promega), 25mM glycylglycine in ddH_2O, pH 7,8 (stored at –20°C)
reaction buffer	1:5 ATP solution : luciferase buffer

10µl of the protein lysate (4.8.1) were added to 400µl of reaction buffer in a suitable tube and emitted light was measured with a Lumat LB9507 (Berthold) Luminometer after automatic injection of 50µl luciferin (measurement period: 10s). Luciferase activity was measured in relative light units (rlu).

4.8.3 BCA assay for determining protein concentration of samples

Results from the luciferase assay were normalised by relating them to the total protein content of the sample. Amounts of protein were determined by the BCA assay.

BCA A solution	1g bicinchoninic acid 1,71g Na_2CO_3 0,95g $NaHCO_3$ 0,16g Na_2Tartrat 100ml ddH_2O pH set to 11,25 with 10N NaOH (stored at 4°C)
BCA B solution	4g $CuSO_4$ x 5 H_2O/100ml dissolved in ddH_2O (stored at 4°C)
Reaction solution	98% BCA A, 2% BCA B, freshly prepared
Protein standard	5mg/ml lysozyme in ddH_2O (stored at –20°C)

The first row of wells of a microtiter plate for optical tests was filled with 190µl of the reaction solution each. 100µl of the solution were added to all other wells. 10µl of Tris-HCL (pH 7,6) were pipetted into the first well of the first column to determine the background signal. 10µl of the protein standard were added to the first well of the second column to establish a calibration curve. The remaining wells of the first row were topped up with 10µl of each protein lysate. Contents of the first row were mixed by repeated pipetting using a multichannel pipette. 100µl of the first row were then transferred to the second row and mixed again. This was repeated to the last row of the dish and finally the excess of 100µl/column was discarded. The plate was incubated for 30-60min at 37°C and absorption at 595nm was measured by a Multiskan EX reader (Thermo Electron Corporation). Values from the first column were automatically subtracted. Protein concentrations were calculated on the basis of the calibration curve.

4.9 Mouse strains

C57BL/6	Harlan
Balb/C	Harlan
129/Ola	Since transgenic mice were derived of IB10 mES cells injected into Balb/C blastocysts, they were of a mixed 129/Ola/Balb/C background
CD1	Harlan; females were used as foster animals for blastocyst transfers
Rag-2/interleukin 2 gamma chain double KO mice (immunodeficient mice lacking B, T and NK cells; first described by Mazurier *et al*., 1999)	Double KO mice were generated by mating Rag-2 deficient animals (Shinkai *et al*., 1992) to the interleukin 2 receptor gamma chain KO strain (DiSanto *et al*., 1995); Double KO mice were of C57BL/6 background

4.9.1 Breeding and keeping of mice

Mice were bred and kept according to the common obligations in the central animal facility of the Helmholtz Centre for Infection Research. When needed, animals were administered with 2mg/ml dox via the drinking water or with 2mg/100µl by gavage.

4.9.2 Induction of teratomas in mice

For teratoma induction, mES cells were harvested with TEP, counted and washed with PBS. After centrifugation (1000rpm, 5min), the pellet was resuspended in fresh PBS and kept on ice until 100-200µl containing 0,5-2x 10^6 were injected subcutaneously into the sides of anaesthetised Rag-2/interleukin 2 gamma chain double KO mice. These mice are immunodeficient and unable to reject the cells. Some, but not all injected mice developed tumors after 3-4 weeks. A good qualitiy of injected cells is critical for teratoma formation efficiency. Tumor bearing animals were subjected to dox treatment and imaged by *in vivo* BLI as described (figure 16). Of note, some factors limit the quantification of luciferase signals from teratomas: 1) individual tumors display different growth rates, 2) images of the same tumor on different days are not directly comparable since increase in size also augments luciferase signals, and 3) luciferin might penetrate individual teratomas with different rates. Still, analysis of teratomas derived from transgenic mES cells is a valuable tool to qualitatively evaluate transgene expression after differentiation since behaviour of transgenic mice seems to be faithfully mimicked.

4.10 Bioluminescence imaging with Xenogen IVIS 200

The *in vivo* imaging technology (Xenogen/Caliper) provides means to repeatedly monitor luciferase activity in the same animal without the need to sacrifice it. Also, a large number of cell clones can be quickly screened for reporter gene expression. First, a grey-scale image of the sample in the light tight chamber is automatically taken. Then, photons are collected by a sensitive CCD (charge-coupled device) camera and the signals are overlayed to the grey-scale image. Parameters like aperture of the lense, exposition time and binning allow to regulate sensitivity and resolution of the image. The field of view may be adjusted to image up to 5 mice in one group. Settings for each immage are given in the respective figure legend. Analyses of images were performed with the Living image 2.60.1 (Igor Pro 4.09A) computer programme.

4.10.1 ... *in vivo*

Animals were first anaesthetised in the induction chamber by 2-2,5% isoflurane (Abbot). In the case of black coat colour, body parts that needed to be imaged were depilated by treatment with commercial depilatory cream since dark fur absorbs a high percentage of the light signal. Mice were then injected i.p. with 100µl of lucifern (30mg/ml in PBS, Synchem OHG) and put on the heated (37°) platform in the acquisition chamber. Anaesthesia was maintained by constant administration of isoflurane via nose cones and images were taken.
For imaging of extracted organs from transgenic mice, animals were injected with luciferin and sacrificed by cercvical dislocation after 3-4min. Organs were removed and immediately imaged in a petri dish.

4.10.2 ... *in vitro*

mES cells tagged with pTagPAC/ΔNPT were grown to near confluence on gelatinised microtiter plates in the presence of dox for 3d. For imaging, wells were washed with PBS and 50µl of 0,1mM luciferin dissolved in reaction buffer as described in 4.8.2 were added to each well. Plates were imaged 2-3min later.

5 Vectors and Oligonucleotides

Note that two vectors were termed pAutoTar* since they carry the same array of features. However, they differ in the reading frame that reconstitutes/maintains *npt* function upon targeting. pAutoTar(1) (pLuc3ATG, # 2593) was used to target pTagTK/NPT tagged sites. pAutoTar(2)* (pEMLuc3rTA2E, # 2741) was used to exchange the pTagPAC/ΔNPT parental locus. The ROSA26 locus in IB10 ES cells had been modified by integration of pROSAantiluc via HR by U. Sandhu. The cell clone harbouring the ROSA26 locus tagged with heterologous FRT sites was called "C7Rosaluc". RMCE in C7Rosaluc yielded the exchanged subclones ROSAautoluc (pAutoTar(2)), ROSAEMLuc (pEM-luc3), and ROSA**rtTA**autoLuc (pEM-rTA2Luc3rTA2). These targetings were performed by S. Bantner.

5.1 Applied vectors

F3PGKbgF (# 1926)	J. Seibler, 1999
flpepuro (pCAGGSFlpeIRESpuro) (# 1825)	F. Stewart, 1999, formerly EMBL Heidelberg
pHoxB4-2A (SF91- eGFP2AHOX+wPRE) (# 3280)	Pilat *et al.*, 2005
pAutoTar(2)* (pEMLuc3rTA2E) (# 2741)	unpublished, S. Herrmann, 2005
pDsredExpress-MCS (# 2566)	unpublished, A. Oumard, 2004
pEGFPN1-MCS (# 1957)	unpublished, A. Baer, 2000
pEM-luc3 (# 3032)	unpublished, S. Bantner, 2006
pEM-rTA2Luc3rTA2 (# 2844)	unpublished, S. Bantner, 2006
pTagTK/NPT (Luc3rTA2E) (# 2375)	C. Wodarczyk, 2003
pSVGP1 (# 604)	M. Wirth, 1990; the expression plasmid contains the MOV gagpol region with the 5' untranslated leader up to AatII under control of the SV40 early promoter and late pA signals
pVSV-G (# 2855)	Invitrogen; CMV driven VSV-G; helper plasmid for lentivirus production

5.2 Cloned vectors

pAutoTar(1)* (pLuc3ATG) (2593)	pTagTK/NPT (# 2375) was cut with *SalI* and *BstZ17I* (loss of EMCV IRES and tk/npt). TGALTA2E (# 2384) was cut with *NheI*, filled in by Klenow, cut with *SalI*. The 757bp fragment from TGALTA2E was ligated into the backbone of pTagTK/NPT
pCflpe (# 2703)	flpepuro (# 1825) was cut with *AvrII* (loss of *pac* and most of the EMCV IRES). The vector was religated.
pFCL3EPF5 (# 2719)	pCL3EPF5 (# 2718) was cut with *SpeI*, blunt ends created by Klenow, cut with *PvuI*. pL3P5 (# 2701) was cut with *HpaI* and *PvuI*. The 1,0 kb fragment containing the FRT wt site was ligated into the pCL3EPF5 backbone.
pFCL3ETKNF5 (# 2727)	pFCL3EPF5 (# 2719) was cut with *PvuI* and *PflMI* (loss of *pac*, SV40pA, FRT5, oriC and part of *amp*). pAutoTar(1) (# 2593) was cut with *PvuI* and *PflMI* and the 1,5 kb fragment containing FRT5, oriC and the complementing part of *amp* was ligated into the backbone.
pTagPAC/ΔNPT (pL3rtTAdneo) (# 3210)	pTagTK/NPT (# 2375) was cut with *PflMI* and *PvuII* for generating the 7,4 kb backbone (removal of part of the EMCV IRES; *tk* gene; FRT-5 site and part of *npt* including loxP). pROSAantiluc (# 2684) was cut with *PflMI* and *PvuII* and the 1,9 kb fragment containing part of EMCV IRES, *pac*, SV40 pA, FRT5 and part of Δ*npt* was ligated into the backbone.

5.3 Oligomeres

Primers were ordered from MWG. For amplification of fragments containing the FRT sites from clone 6.2 genomic DNA, primer pairs P1/P3 (FRT5) and 5'AmpProm/ L3end (FRT wt) were used.

P1 (EMCV2, P2451)	5' GCCACGTTGTGAGTTGGATA 3'
P2 (JunPA12, P834)	5' GAGGAAATTGCATCGCATTGT 3'
P3 (neorev2, P1523)	5' GTCATAGCCGAATAGCCTCTCC 3'
5'AmpProm (P1362)	5' CGACACGGAAATGTTGAATA 3'
L3end (P2374)	5' GCGGAAAGATCGCCGTG 3'

6 Abbreviations

5-aza	5-azacytidine
A	adenosine
amp	ampicillin resistance gene
ATCC	American Type Culture Collection
ATG	translational start
ATP	adenosine triphosphate
attB	recognition target of ΦC31
attP	recognition target of ΦC31
attL	recognition target of ΦC31, after recombination
attR	recognition target of ΦC31, after recombination
BAC	bacterial artificial chromosome
bp	base pair(s)
BLI	bioluminescent imaging
BSA	bovine serum albumine
C	cytosine
CAGGS = CAG	Cytomegalo virus/chicken β-actin promoter
CHO cells	chinese hamster ovary cell line
CID	chemically induced dimerisation
CMV	Cytomegalo virus
cp.	compare
Cre	Cre recombinase (cyclisation recombination)
CreER(T2)	Cre recombinase fused to mutated oestrogen receptor binding domain
d	day(s)
ddH_2O	double distilled H_2O
Δ*npt*	5’ truncated *npt* gene
DMEM	Dulbecco’s modified Eagle’s medium
DMSO	dimethylsulfoxide
DNA	deoxyribo nucleic acid
dNTP	deoxyribo nucleotide triphosphate
dox	doxycycline
Dre	Dre recombinase
DSMZ	German Collection of Microorganisms and Cell Cultures
dsRed	red fluorescent protein
DTT	dithiothreitol
EB	embryoid body
E. coli	*Escherichia coli*
EDTA	ethylenediamine tetraacetic acid
eGFP	enhanced green fluorescent protein
EMCV	Encephalomyocarditis virus
ES-HCs	hematopoietic progenitors derived *in vitro* from mES cells
FACS	fluorescence activated cell sorting
FCS	fetal calf serum
FISH	Fluorescence *in situ* hybridisation
Flp	Flp recombinase
Flpe	enhanced Flp recombinase
FlpeER(T2)	Flpe recombinase fused to mutated oestrogen receptor binding domain
FRT	Flp recombinase recognition target
G	guanine

G418	aminoglycoside-2'-deoxystreptine (gentamycin derivative)
GCV	ganciclovir
GFP	green fluorescent protein
GOI	gene of interest
h	hour(s)
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HeLa	human cervix carcinom cell line
HEPES	N-2-hydroxyethylpiperazine-N'-ethansulfonic acid
HMW	high molecular weight (genomic) DNA
HPRT	hypoxanthine-guanine phosphoribosyl transferase
HR	homologous recombination
HSC	hematopoietic stem cells
Hsp90	heat shock protein 90 complex
HSV	Herpes simplex virus
IGTC	International Gene Trap Consortium
IMDM	Iscove's modified Dulbecco's medium
i.p.	intraperitoneal
IPTG	isopropyl-beta-D-thiogalactopyranoside
IRES	internal ribosomal entry site
kb	1000 base(s)
kbp	1000 base pair(s)
KO	knock out
KRAB	Krüppel-associated box repressor domain
kV	1000 volt(s)
l	liter(s)
LBD	ligand binding domain
LacI	Lac repressor
lacO	operator sequence of the *lac* system
lacZ	β-galactosidase gene
LIF	leukaemia inhibitory factor
loxP	locus of crossover (Cre recognition target)
MEF	murine embryonic fibroblasts
mES cells	murine embryonic stem cells
min	minute(s)
ml	milliliter(s)
moi	multiplicity of infection
mRNA	messenger RNA
MTG	monothioglycerol
µF	micro Farad
npt	neomycin phosphotransferase gene
nls	nuclear localisation signal
Ω	ohm(s)
ORF	open reading frame
ori	origin of replication
pA	poly adenylation signal
pac	puromycin acetyl transferase gene
P_{bi}	bidirectional, tet dependent promoter
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PGK	phosphoglycerate kinase

ΦC31	ΦC31 recombinase
RAG-1	recombination activating gene 1
RAG-2	recombination activating gene 2
RI	random integration
rlu	relative light units
RMCE	recombinase mediated cassette exchange
RNA	ribo nucleic acid
RNAi	RNA interference
rpm	rounds per minute
RT	recognition target
rtTA	reverse transactivator of the tet system
rtTA2(S)-M2	reverse transactivator of the tet system
rtTS	reverse transrepressor of the tet system
s	second(s)
SBD	steroid binding domain
SDS	sodium dodecyl sulfate
shRNA	small hairpin RNA
siRNA	small interfering RNA
SSC	sodium chloride/ sodium citrate
SSR	site specific recombinases
T	thymidine
TEP	trypsin-EDTA
tet	tetracycline
tetO	operator sequence of the tet system
Tet_{ON}	tet dependent expression system
Tet_{OFF}	tet dependent expression system
TetR	tet system repressor domain
Tris	2-amino-2-hydroxymethyl-1,3-propanediol
tTA	transcactivator of the tet system
tTA2	transcactivator of the tet system
tTA4	transcactivator of the tet system
tTS	transrepressor of the tet system
tk	thymidine kinase gene
tk/npt	fusion of *tk* and *npt*
TSA	trychostatin A
U	unit
V	volt(s)
VSV	Vesicular stomatitis virus
v/v	volume/volume; percent by volume
wt	wild type
w/v	weight/volume; percent by weight

7 References

Akagi, K., Kanai, M., Saya, H., Kozu, T. & Berns, A. A novel tetracycline-dependent transactivator with E2F4 transcriptional activation domain. *Nucleic Acids Res* **29**, E23 (2001).

Amin, A., Roca, H., Luetke, K. & Sadowski, P.D. Synapsis, strand scission, and strand exchange induced by the FLP recombinase: analysis with half-FRT sites. *Mol Cell Biol* **11**, 4497-508 (1991).

Andreas, S., Schwenk, F., Kuter-Luks, B., Faust, N. & Kuhn, R. Enhanced efficiency through nuclear localization signal fusion on phage PhiC31-integrase: activity comparison with Cre and FLPe recombinase in mammalian cells. *Nucleic Acids Res* **30**, 2299-306 (2002).

Araki, K., Imaizumi, T., Okuyama, K., Oike, Y. & Yamamura, K. Efficiency of recombination by Cre transient expression in embryonic stem cells: comparison of various promoters. *J Biochem (Tokyo)* **122**, 977-82 (1997).

Austin, C.P. *et al*. The knockout mouse project. *Nat Genet* **36**, 921-4 (2004).

Auwerx, J. *et al*. The European dimension for the mouse genome mutagenesis program. *Nat Genet* **36**, 925-7 (2004).

Awatramani, R., Soriano, P., Rodriguez, C., Mai, J.J. & Dymecki, S.M. Cryptic boundaries in roof plate and choroid plexus identified by intersectional gene activation. *Nat Genet* **35**, 70-5 (2003).

Baer, A. & Bode, J. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Curr Opin Biotechnol* **12**, 473-80 (2001).

Baron, U., Freundlieb, S., Gossen, M. & Bujard, H. Co-regulation of two gene activities by tetracycline via a bidirectional promoter. *Nucleic Acids Res* **23**, 3605-6 (1995).

Baron, U., Gossen, M. & Bujard, H. Tetracycline-controlled transcription in eukaryotes: novel transactivators with graded transactivation potential. *Nucleic Acids Res* **25**, 2723-9 (1997).

Baron, U. *et al*. Generation of conditional mutants in higher eukaryotes by switching between the expression of two genes. *Proc Natl Acad Sci U S A* **96**, 1013-8 (1999).

Becskei, A., Seraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *Embo J* **20**, 2528-35 (2001).

Belteki, G. *et al*. Conditional and inducible transgene expression in mice through the combinatorial use of Cre-mediated recombination and tetracycline induction. *Nucleic Acids Res* **33**, e51 (2005).

Bender, B. *et al*. Position independent and copy-number-related expression of the bovine neonatal Fc receptor alpha-chain in transgenic mice carrying a 102 kb BAC genomic fragment. *Transgenic Res* **16**, 613-27 (2007).

Biggar, S.R. & Crabtree, G.R. Cell signaling can direct either binary or graded transcriptional responses. *Embo J* **20**, 3167-76 (2001).

Bode, J. *et al*. The transgeneticist's toolbox: novel methods for the targeted modification of eukaryotic genomes. *Biol Chem* **381**, 801-13 (2000).

Bonner, A.E., Wang, Y. & You, M. Gene expression profiling of mouse teratocarcinomas uncovers epigenetic changes associated with the transformation of mouse embryonic stem cells. *Neoplasia* **6**, 490-502 (2004).

Bornkamm, G.W. *et al*. Stringent doxycycline-dependent control of gene activities using an episomal one-vector system. *Nucleic Acids Res* **33**, e137 (2005).

Boshart, M. *et al*. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. *Cell* **41**, 521-30 (1985).

Branda, C.S. & Dymecki, S.M. Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. *Dev Cell* **6**, 7-28 (2004).

Braun, R.E. *et al*. Infertility in male transgenic mice: disruption of sperm development by HSV-tk expression in postmeiotic germ cells. *Biol Reprod* **43**, 684-93 (1990).

Buchholz, F., Angrand, P.O. & Stewart, A.F. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat Biotechnol* **16**, 657-62 (1998).

Buchholz, F., Ringrose, L., Angrand, P.O., Rossi, F. & Stewart, A.F. Different thermostabilities of FLP and Cre recombinases: implications for applied site-specific recombination. *Nucleic Acids Res* **24**, 4256-62 (1996).

Candau, R. *et al*. Identification of human proteins functionally conserved with the yeast putative adaptors ADA2 and GCN5. *Mol Cell Biol* **16**, 593-602 (1996).

Cesari, F. *et al*. Elk-1 knock-out mice engineered by Flp recombinase-mediated cassette exchange. *Genesis* **38**, 87-92 (2004).

Chtarto, A. *et al*. Tetracycline-inducible transgene expression mediated by a single AAV vector. *Gene Ther* **10**, 84-94 (2003).

Cobellis, G. *et al*. Tagging genes with cassette-exchange sites. *Nucleic Acids Res* **33**, e44 (2005).

Cronin, C.A., Gluba, W. & Scrable, H. The lac operator-repressor system is functional in the mouse. *Genes Dev* **15**, 1506-17 (2001).

Cronin, C.A., Ryan, A.B., Talley, E.M. & Scrable, H. Tyrosinase expression during neuroblast divisions affects later pathfinding by retinal ganglion cells. *J Neurosci* **23**, 11692-7 (2003).

Damke, H., Gossen, M., Freundlieb, S., Bujard, H. & Schmid, S.L. Tightly regulated and inducible expression of dominant interfering dynamin mutant in stably transformed HeLa cells. *Methods Enzymol* **257**, 209-20 (1995).

Deuschle, U., Meyer, W.K. & Thiesen, H.J. Tetracycline-reversible silencing of eukaryotic promoters. *Mol Cell Biol* **15**, 1907-14 (1995).

Dickins, R.A. *et al*. Tissue-specific and reversible RNA interference in transgenic mice. *Nat Genet* **39**, 914-21 (2007).

DiSanto, J.P., Muller, W., Guy-Grand, D., Fischer, A. & Rajewsky, K. Lymphoid development in mice with a targeted deletion of the interleukin 2 receptor gamma chain. *Proc Natl Acad Sci U S A* **92**, 377-81 (1995).

Downing, G.J. & Battey, J.F., Jr. Technical assessment of the first 20 years of research using mouse embryonic stem cell lines. *Stem Cells* **22**, 1168-80 (2004).

Ernst, E. *et al*. Generation of inducible hepatitis C virus transgenic mouse lines. *J Med Virol* **79**, 1103-12 (2007).

Evans, M.J. & Kaufman, M.H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-6 (1981).

Farley, F.W., Soriano, P., Steffen, L.S. & Dymecki, S.M. Widespread recombinase expression using FLPeR (flipper) mice. *Genesis* **28**, 106-10 (2000).

Feil, R., Wagner, J., Metzger, D. & Chambon, P. Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. *Biochem Biophys Res Commun* **237**, 752-7 (1997).

Feng, G. *et al*. Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron* **28**, 41-51 (2000).

Fiering, S., Whitelaw, E. & Martin, D.I. To be or not to be active: the stochastic nature of enhancer action. *Bioessays* **22**, 381-7 (2000).

Figge, J., Wright, C., Collins, C.J., Roberts, T.M. & Livingston, D.M. Stringent regulation of stably integrated chloramphenicol acetyl transferase genes by E. coli lac repressor in monkey cells. *Cell* **52**, 713-22 (1988).

Flint, J. & Shenk, T. Viral transactivating proteins. *Annu Rev Genet* **31**, 177-212 (1997).

Forster, K. *et al*. Tetracycline-inducible expression systems with reduced basal activity in mammalian cells. *Nucleic Acids Res* **27**, 708-10 (1999).

Fraser, H.B., Hirsh, A.E., Giaever, G., Kumm, J. & Eisen, M.B. Noise minimization in eukaryotic gene expression. *PLoS Biol* **2**, e137 (2004).

Freundlieb, S., Schirra-Muller, C. & Bujard, H. A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. *J Gene Med* **1**, 4-12 (1999).

Friedrich, G. & Soriano, P. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* **5**, 1513-23 (1991).

Frohberg, C., Heins, L. & Gatz, C. Characterization of the interaction of plant transcription factors using a bacterial repressor protein. *Proc Natl Acad Sci U S A* **88**, 10470-4 (1991).

Furth, P.A. *et al*. Temporal control of gene expression in transgenic mice by a tetracycline-responsive promoter. *Proc Natl Acad Sci U S A* **91**, 9302-6 (1994).

Fussenegger, M. The impact of mammalian gene regulation concepts on functional genomic research, metabolic engineering, and advanced gene therapies. *Biotechnol Prog* **17**, 1-51 (2001).

Fussenegger, M. *et al*. Streptogramin-based gene regulation systems for mammalian cells. *Nat Biotechnol* **18**, 1203-8 (2000).

Galla, M., Will, E., Kraunus, J., Chen, L. & Baum, C. Retroviral pseudotransduction for targeted cell manipulation. *Mol Cell* **16**, 309-15 (2004).

Gallia, G.L. & Khalili, K. Evaluation of an autoregulatory tetracycline regulated system. *Oncogene* **16**, 1879-84 (1998).

Gao, Q. *et al*. Telomeric Transgenes are Silenced in Adult Mouse Tissues and Embryo Fibroblasts, but are Expressed in Embryonic Stem Cells. *Stem Cells* (2007).

Garrick, D., Fiering, S., Martin, D.I. & Whitelaw, E. Repeat-induced gene silencing in mammals. *Nat Genet* **18**, 56-9 (1998).

Garrick, D., Sutherland, H., Robertson, G. & Whitelaw, E. Variegated expression of a globin transgene correlates with chromatin accessibility but not methylation status. *Nucleic Acids Res* **24**, 4902-9 (1996).

Gatz, C., Kaiser, A. & Wendenburg, R. Regulation of a modified CaMV 35S promoter by the Tn10-encoded Tet repressor in transgenic tobacco. *Mol Gen Genet* **227**, 229-37 (1991).

Gatz, C. & Quail, P.H. Tn10-encoded tet repressor can regulate an operator-containing plant promoter. *Proc Natl Acad Sci U S A* **85**, 1394-7 (1988).

Giel-Moloney, M., Krause, D.S., Chen, G., Van Etten, R.A. & Leiter, A.B. Ubiquitous and uniform in vivo fluorescence in ROSA26-EGFP BAC transgenic mice. *Genesis* **45**, 83-9 (2007).

Gill, G. & Ptashne, M. Negative effect of the transcriptional activator GAL4. *Nature* **334**, 721-4 (1988).

Ginsberg, D. *et al*. E2F-4, a new member of the E2F transcription factor family, interacts with p107. *Genes Dev* **8**, 2665-79 (1994).

Giraldo, P. & Montoliu, L. Size matters: use of YACs, BACs and PACs in transgenic animals. *Transgenic Res* **10**, 83-103 (2001).

Gong, S. *et al*. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* **425**, 917-25 (2003).

Goodrich, J.A., Hoey, T., Thut, C.J., Admon, A. & Tjian, R. Drosophila TAFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. *Cell* **75**, 519-30 (1993).

Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A* **89**, 5547-51 (1992).

Gossen, M. *et al*. Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**, 1766-9 (1995).

Gould, D.J. *et al*. A novel doxycycline inducible autoregulatory plasmid which displays "on"/"off" regulation suited to gene therapy applications. *Gene Ther* **7**, 2061-70 (2000).

Gould, D.J. & Chernajovsky, Y. Endogenous GATA factors bind the core sequence of the tetO and influence gene regulation with the tetracycline system. *Mol Ther* **10**, 127-38 (2004).

Groth, A.C., Olivares, E.C., Thyagarajan, B. & Calos, M.P. A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci U S A* **97**, 5995-6000 (2000).

Hayakawa, T., Yusa, K., Kouno, M., Takeda, J. & Horie, K. Bloom's syndrome gene-deficient phenotype in mouse primary cells induced by a modified tetracycline-controlled trans-silencer. *Gene* **369**, 80-9 (2006).

Hennighausen, L., Wall, R.J., Tillmann, U., Li, M. & Furth, P.A. Conditional gene expression in secretory tissues and skin of transgenic mice using the MMTV-LTR and the tetracycline responsive system. *J Cell Biochem* **59**, 463-72 (1995).

Hayes, S. & O'Hare, P. Mapping of a major surface-exposed site in herpes simplex virus protein Vmw65 to a region of direct interaction in a transcription complex assembly. *J Virol* **67**, 852-62 (1993).

Heo, J. *et al*. Hepatic precursors derived from murine embryonic stem cells contribute to regeneration of injured liver. *Hepatology* **44**, 1478-86 (2006).

Hillen, W. & Berens, C. Mechanisms underlying expression of Tn10 encoded tetracycline resistance. *Annu Rev Microbiol* **48**, 345-69 (1994).

Hinrichs, W. *et al*. Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. *Science* **264**, 418-20 (1994).

Hitz, C., Wurst, W. & Kuhn, R. Conditional brain-specific knockdown of MAPK using Cre/loxP regulated RNA interference. *Nucleic Acids Res* **35**, e90 (2007).

Hoess, R., Wierzbicki, A. & Abremski, K. Formation of small circular DNA molecules via an in vitro site-specific recombination system. *Gene* **40**, 325-9 (1985a).

Hoess, R.H. & Abremski, K. Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. *J Mol Biol* **181**, 351-62 (1985b).

Hoess, R.H., Wierzbicki, A. & Abremski, K. The role of the loxP spacer region in P1 site-specific recombination. *Nucleic Acids Res* **14**, 2287-300 (1986).

Hofmann, A., Nolan, G.P. & Blau, H.M. Rapid retroviral delivery of tetracycline-inducible genes in a single autoregulatory cassette. *Proc Natl Acad Sci U S A* **93**, 5185-90 (1996).

Hong, H.K. *et al*. Inducible and reversible Clock gene expression in brain using the tTA system for the study of circadian behavior. *PLoS Genet* **3**, e33 (2007).

Hooper, M., Hardy, K., Handyside, A., Hunter, S. & Monk, M. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* **326**, 292-5 (1987).

Howe, J.R., Skryabin, B.V., Belcher, S.M., Zerillo, C.A. & Schmauss, C. The responsiveness of a tetracycline-sensitive expression system differs in different cell lines. *J Biol Chem* **270**, 14168-74 (1995).

Hu, M.C. & Davidson, N. The inducible lac operator-repressor system is functional in mammalian cells. *Cell* **48**, 555-66 (1987).

Hunter, N.L., Awatramani, R.B., Farley, F.W. & Dymecki, S.M. Ligand-activated Flpe for temporally regulated gene modifications. *Genesis* **41**, 99-109 (2005).

Ingles, C.J., Shales, M., Cress, W.D., Triezenberg, S.J. & Greenblatt, J. Reduced binding of TFIID to transcriptionally compromised mutants of VP16. *Nature* **351**, 588-90 (1991).

Isaacs, F.J., Hasty, J., Cantor, C.R. & Collins, J.J. Prediction and measurement of an autoregulatory genetic module. *Proc Natl Acad Sci U S A* **100**, 7714-9 (2003).

Jonnalagadda, V.S., Matsuguchi, T. & Engelward, B.P. Interstrand crosslink-induced homologous recombination carries an increased risk of deletions and insertions. *DNA Repair (Amst)* **4**, 594-605 (2005).

Kaern, M., Elston, T.C., Blake, W.J. & Collins, J.J. Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Genet* **6**, 451-64 (2005).

Kamper, M.R., Gohla, G. & Schluter, G. A novel positive tetracycline-dependent transactivator (rtTA) variant with reduced background activity and enhanced activation potential. *FEBS Lett* **517**, 115-20 (2002).

Keller, G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev* **19**, 1129-55 (2005).

Kistner, A. *et al*. Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. *Proc Natl Acad Sci U S A* **93**, 10933-8 (1996).

Kramer, B.P., Weber, W. & Fussenegger, M. Artificial regulatory networks and cascades for discrete multilevel transgene control in mammalian cells. *Biotechnol Bioeng* **83**, 810-20 (2003).

Kringstein, A.M., Rossi, F.M., Hofmann, A. & Blau, H.M. Graded transcriptional response to different concentrations of a single transactivator. *Proc Natl Acad Sci U S A* **95**, 13670-5 (1998).

Krishnan, M. *et al*. Effects of epigenetic modulation on reporter gene expression: implications for stem cell imaging. *Faseb J* **20**, 106-8 (2006).

Krueger, C., Berens, C., Schmidt, A., Schnappinger, D. & Hillen, W. Single-chain Tet transregulators. *Nucleic Acids Res* **31**, 3050-6 (2003).

Kuhnel, F. *et al*. Doxycycline regulation in a single retroviral vector by an autoregulatory loop facilitates controlled gene expression in liver cells. *Nucleic Acids Res* **32**, e30 (2004).

Kyba, M., Perlingeiro, R.C. & Daley, G.Q. HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell* **109**, 29-37 (2002).

Langa, F. *et al*. Teratocarcinomas induced by embryonic stem (ES) cells lacking vimentin: an approach to study the role of vimentin in tumorigenesis. *J Cell Sci* **113 Pt 19**, 3463-72 (2000).

Lee, P. *et al*. Conditional lineage ablation to model human diseases. *Proc Natl Acad Sci U S A* **95**, 11371-6 (1998).

Leuchtenberger, S., Perz, A., Gatz, C. & Bartsch, J.W. Conditional cell ablation by stringent tetracycline-dependent regulation of barnase in mammalian cells. *Nucleic Acids Res* **29**, E76 (2001).

Lin, Y.S., Ha, I., Maldonado, E., Reinberg, D. & Green, M.R. Binding of general transcription factor TFIIB to an acidic activating region. *Nature* **353**, 569-71 (1991).

Liu, J., Jeppesen, I., Nielsen, K. & Jensen, T.G. Phi c31 integrase induces chromosomal aberrations in primary human fibroblasts. *Gene Ther* **13**, 1188-90 (2006a).

Liu, K. *et al*. Recombinase-mediated cassette exchange to rapidly and efficiently generate mice with human cardiac sodium channels. *Genesis* **44**, 556-64 (2006b).

Loonstra, A. *et al*. Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells. *Proc Natl Acad Sci U S A* **98**, 9209-14 (2001).

Mader, S., Chambon, P. & White, J.H. Defining a minimal estrogen receptor DNA binding domain. *Nucleic Acids Res* **21**, 1125-32 (1993).

Margolin, J.F. *et al*. Kruppel-associated boxes are potent transcriptional repression domains. *Proc Natl Acad Sci U S A* **91**, 4509-13 (1994).

Markusic, D., Oude-Elferink, R., Das, A.T., Berkhout, B. & Seppen, J. Comparison of single regulated lentiviral vectors with rtTA expression driven by an autoregulatory loop or a constitutive promoter. *Nucleic Acids Res* **33**, e63 (2005).

Martin, G.R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* **78**, 7634-8 (1981).

Masui, S. *et al*. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* **9**, 625-35 (2007).

Masui, S. *et al*. An efficient system to establish multiple embryonic stem cell lines carrying an inducible expression unit. *Nucleic Acids Res* **33**, e43 (2005).

Matthaei, K.I. Genetically manipulated mice: a powerful tool with unsuspected caveats. *J Physiol* **582**, 481-8 (2007).

Matthess, Y. *et al*. Conditional inhibition of cancer cell proliferation by tetracycline-responsive, H1 promoter-driven silencing of PLK1. *Oncogene* **24**, 2973-80 (2005).

May, T., Entwicklung eines transkriptionell regulierten Systems zur konditionalen Immortalisierung. Dissertation, Technical University Braunschweig (2004).

May, T., Hauser, H. & Wirth, D. Current status of transcriptional regulation systems. *Cytotechnology* **50**, 109-119 (2006).

Mazurier, F. *et al*. A novel immunodeficient mouse model--RAG2 x common cytokine receptor gamma chain double mutants--requiring exogenous cytokine administration for human hematopoietic stem cell engraftment. *J Interferon Cytokine Res* **19**, 533-41 (1999).

Migliaccio, A.R. *et al*. Stable and unstable transgene integration sites in the human genome: extinction of the Green Fluorescent Protein transgene in K562 cells. *Gene* **256**, 197-214 (2000).

Muller, G. *et al*. Characterization of non-inducible Tet repressor mutants suggests conformational changes necessary for induction. *Nat Struct Biol* **2**, 693-703 (1995).

Mullick, A. *et al*. The cumate gene-switch: a system for regulated expression in mammalian cells. *BMC Biotechnol* **6**, 43 (2006).

Neddermann, P. *et al*. A novel, inducible, eukaryotic gene expression system based on the quorum-sensing transcription factor TraR. *EMBO Rep* **4**, 159-65 (2003).

Niwa, H., Miyazaki, J. & Smith, A.G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**, 372-6 (2000).

Olivares, E.C. *et al*. Site-specific genomic integration produces therapeutic Factor IX levels in mice. *Nat Biotechnol* **20**, 1124-8 (2002).

Orth, P. *et al*. Conformational changes of the Tet repressor induced by tetracycline trapping. *J Mol Biol* **279**, 439-47 (1998).

Ortiz-Urda, S. *et al*. Stable nonviral genetic correction of inherited human skin disease. *Nat Med* **8**, 1166-70 (2002).

Pankiewicz, R., Karlen, Y., Imhof, M.O. & Mermod, N. Reversal of the silencing of tetracycline-controlled genes requires the coordinate action of distinctly acting transcription factors. *J Gene Med* **7**, 117-32 (2005).

Peitz, M., Pfannkuche, K., Rajewsky, K. & Edenhofer, F. Ability of the hydrophobic FGF and basic TAT peptides to promote cellular uptake of recombinant Cre recombinase: a tool for efficient genetic engineering of mammalian genomes. *Proc Natl Acad Sci U S A* **99**, 4489-94 (2002).

Picard, D., Salser, S.J. & Yamamoto, K.R. A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor. *Cell* **54**, 1073-80 (1988).

Pilat, S. *et al*. HOXB4 enforces equivalent fates of ES-cell-derived and adult hematopoietic cells. *Proc Natl Acad Sci U S A* **102**, 12101-6 (2005).

Puttini, S. *et al*. Development of a targeted transgenesis strategy in highly differentiated cells: a powerful tool for functional genomic analysis. *J Biotechnol* **116**, 145-51 (2005).

Postle, K., Nguyen, T.T. & Bertrand, K.P. Nucleotide sequence of the repressor gene of the TN10 tetracycline resistance determinant. *Nucleic Acids Res* **12**, 4849-63 (1984).

Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol* **4**, e309 (2006).

Raymond, C.S. & Soriano, P. High-efficiency FLP and PhiC31 site-specific recombination in mammalian cells. *PLoS ONE* **2**, e162 (2007).

Reid, L.H., Shesely, E.G., Kim, H.S. & Smithies, O. Cotransformation and gene targeting in mouse embryonic stem cells. *Mol Cell Biol* **11**, 2769-77 (1991).

Reppel, M. *et al*. Effect of cardioactive drugs on action potential generation and propagation in embryonic stem cell-derived cardiomyocytes. *Cell Physiol Biochem* **19**, 213-24 (2007).

Robanus-Maandag, E. *et al*. p107 is a suppressor of retinoblastoma development in pRb-deficient mice. *Genes Dev* **12**, 1599-609 (1998).

Robertson, A., Perea, J., Tolmachova, T., Thomas, P.K. & Huxley, C. Effects of mouse strain, position of integration and tetracycline analogue on the tetracycline conditional system in transgenic mice. *Gene* **282**, 65-74 (2002).

Rodriguez, C.I. *et al*. High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. *Nat Genet* **25**, 139-40 (2000).

Rossi, F.M. *et al*. Tetracycline-regulatable factors with distinct dimerization domains allow reversible growth inhibition by p16. *Nat Genet* **20**, 389-93 (1998).

Ryan, A. & Scrable, H. Visualization of the dynamics of gene expression in the living mouse. *Mol Imaging* **3**, 33-42 (2004).

Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. GAL4-VP16 is an unusually potent transcriptional activator. *Nature* **335**, 563-4 (1988).

Sadowski, P.D. The Flp recombinase of the 2-microns plasmid of Saccharomyces cerevisiae. *Prog Nucleic Acid Res Mol Biol* **51**, 53-91 (1995).

Saez, E., No, D., West, A. & Evans, R.M. Inducible gene expression in mammalian cells and transgenic mice. *Curr Opin Biotechnol* **8**, 608-16 (1997).

Sato, N., Nakayama, M. & Arai, K. Fluctuation of chromatin unfolding associated with variation in the level of gene expression. *Genes Cells* **9**, 619-30 (2004).

Sauer, B. & McDermott, J. DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-c1 regions of P1-related phages. *Nucleic Acids Res* **32**, 6086-95 (2004).

Schaft, J., Ashery-Padan, R., van der Hoeven, F., Gruss, P. & Stewart, A.F. Efficient FLP recombination in mouse ES cells and oocytes. *Genesis* **31**, 6-10 (2001).

Schmitz, M.L. & Baeuerle, P.A. The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B. *Embo J* **10**, 3805-17 (1991).

Schmidt-Supprian, M. & Rajewsky, K. Vagaries of conditional gene targeting. *Nat Immunol* **8**, 665-8 (2007).

Schnutgen, F., Stewart, A.F., von Melchner, H. & Anastassiadis, K. Engineering embryonic stem cells with recombinase systems. *Methods Enzymol* **420**, 100-36 (2006).

Schonig, K., Schwenk, F., Rajewsky, K. & Bujard, H. Stringent doxycycline dependent control of CRE recombinase in vivo. *Nucleic Acids Res* **30**, e134 (2002).

Scrable, H. Say when: reversible control of gene expression in the mouse by lac. *Semin Cell Dev Biol* **13**, 109-19 (2002).

Schucht, R. *et al.* A new generation of retroviral producer cells: predictable and stable virus production by Flp-mediated site-specific integration of retroviral vectors. *Mol Ther* **14**, 285-92 (2006).

Seibler, J. Sequenzspezifische Rekombination zur gezielten Manipulation des Mammaliagenoms. Dissertation, Technical University Braunschweig (1999).

Seibler, J. & Bode, J. Double-reciprocal crossover mediated by FLP-recombinase: a concept and an assay. *Biochemistry* **36**, 1740-7 (1997).

Seibler, J. *et al.* Reversible gene knockdown in mice using a tight, inducible shRNA expression system. *Nucleic Acids Res* **35**, e54 (2007).

Senecoff, J.F., Rossmeissl, P.J. & Cox, M.M. DNA recognition by the FLP recombinase of the yeast 2 mu plasmid. A mutational analysis of the FLP binding site. *J Mol Biol* **201**, 405-21 (1988).

Shimshek, D.R. *et al*. Codon-improved Cre recombinase (iCre) expression in the mouse. *Genesis* **32**, 19-26 (2002).

Shin, M.K., Levorse, J.M., Ingram, R.S. & Tilghman, S.M. The temporal requirement for endothelin receptor-B signalling during neural crest development. *Nature* **402**, 496-501 (1999).

Shinkai, Y. *et al*. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* **68**, 855-67 (1992).

Shockett, P., Difilippantonio, M., Hellman, N. & Schatz, D.G. A modified tetracycline-regulated system provides autoregulatory, inducible gene expression in cultured cells and transgenic mice. *Proc Natl Acad Sci U S A* **92**, 6522-6 (1995).

Shockett, P.E., Zhou, S., Hong, X. & Schatz, D.G. Partial reconstitution of V(D)J rearrangement and lymphocyte development in RAG-deficient mice expressing inducible, tetracycline-regulated RAG transgenes. *Mol Immunol* **40**, 813-29 (2004).

Silverman, N., Agapite, J. & Guarente, L. Yeast ADA2 protein binds to the VP16 protein activation domain and activates transcription. *Proc Natl Acad Sci U S A* **91**, 11665-8 (1994).

Smith, M.C. & Thorpe, H.M. Diversity in the serine recombinases. *Mol Microbiol* **44**, 299-307 (2002).

Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. & Kucherlapati, R.S. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317**, 230-4 (1985).

Sparwasser, T. & Eberl, G. BAC to immunology--bacterial artificial chromosome-mediated transgenesis for targeting of immune cells. *Immunology* **121**, 308-13 (2007).

Spencer, D.M., Wandless, T.J., Schreiber, S.L. & Crabtree, G.R. Controlling signal transduction with synthetic ligands. *Science* **262**, 1019-24 (1993).

Sternberg, N., Sauer, B., Hoess, R. & Abremski, K. Bacteriophage P1 cre gene and its regulatory region. Evidence for multiple promoters and for regulation by DNA methylation. *J Mol Biol* **187**, 197-212 (1986).

Strathdee, C.A., McLeod, M.R. & Hall, J.R. Efficient control of tetracycline-responsive gene expression from an autoregulated bi-directional expression vector. *Gene* **229**, 21-9 (1999).

Strathdee, D., Ibbotson, H. & Grant, S.G. Expression of transgenes targeted to the Gt(ROSA)26Sor locus is orientation dependent. *PLoS ONE* **1**, e4 (2006).

Sun, Y., Chen, X. & Xiao, D. Tetracycline-inducible expression systems: new strategies and practices in the transgenic mouse modeling. *Acta Biochim Biophys Sin (Shanghai)* **39**, 235-46 (2007).

Szulc, J., Wiznerowicz, M., Sauvain, M.O., Trono, D. & Aebischer, P. A versatile tool for conditional gene expression and knockdown. *Nat Methods* **3**, 109-16 (2006).

Thomas, K.R., Deng, C. & Capecchi, M.R. High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. *Mol Cell Biol* **12**, 2919-23 (1992).

Thomson, J.A. *et al*. Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145-7 (1998).

Thorpe, H.M. & Smith, M.C. In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc Natl Acad Sci U S A* **95**, 5505-10 (1998).

Toledo, F., Liu, C.W., Lee, C.J. & Wahl, G.M. RMCE-ASAP: a gene targeting method for ES and somatic cells to accelerate phenotype analyses. *Nucleic Acids Res* **34**, e92 (2006).

Triezenberg, S.J., Kingsbury, R.C. & McKnight, S.L. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev* **2**, 718-29 (1988a).

Triezenberg, S.J., LaMarco, K.L. & McKnight, S.L. Evidence of DNA: protein interactions that mediate HSV-1 immediate early gene activation by VP16. *Genes Dev* **2**, 730-42 (1988b).

Tumbar, T., Sudlow, G. & Belmont, A.S. Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. *J Cell Biol* **145**, 1341-54 (1999).

Unsinger, J., Kroger, A., Hauser, H. & Wirth, D. Retroviral vectors for the transduction of autoregulated, bidirectional expression cassettes. *Mol Ther* **4**, 484-9 (2001).

Unsinger, J., Lindenmaier, W., May, T., Hauser, H. & Wirth, D. Stable and strictly controlled expression of LTR-flanked autoregulated expression cassettes upon adenoviral transfer. *Biochem Biophys Res Commun* **319**, 879-87 (2004).

Urlinger, S. *et al*. Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc Natl Acad Sci U S A* **97**, 7963-8 (2000a).

Urlinger, S. *et al*. The p65 domain from NF-kappaB is an efficient human activator in the tetracycline-regulatable gene expression system. *Gene* **247**, 103-10 (2000b).

Wallace, H.A. *et al*. Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. *Cell* **128**, 197-209 (2007).

Weber, W. *et al*. Macrolide-based transgene control in mammalian cells and mice. *Nat Biotechnol* **20**, 901-7 (2002).

Wiese, C., Kania, G., Rolletschek, A., Blyszczuk, P. & Wobus, A.M. Pluripotency: capacity for in vitro differentiation of undifferentiated embryonic stem cells. *Methods Mol Biol* **325**, 181-205 (2006).

Wirth, D. *et al*. Road to precision: recombinase-based targeting technologies for genome engineering. *Curr Opin Biotechnol* (2007).

Wodarczyk, C., Etablierung von ES-Zelllinien zur vorhersagbaren autoregulierten Expression von Transgenen in Mäusen. Dissertation, Technical University Braunschweig (2003).

Wong, E.T. *et al*. Reproducible doxycycline-inducible transgene expression at specific loci generated by Cre-recombinase mediated cassette exchange. *Nucleic Acids Res* **33**, e147 (2005).

Ying, Q.L., Nichols, J., Chambers, I. & Smith, A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115**, 281-92 (2003).

Yoshida, Y. & Hamada, H. Adenovirus-mediated inducible gene expression through tetracycline-controllable transactivator with nuclear localization signal. *Biochem Biophys Res Commun* **230**, 426-30 (1997).

Yu, H.M., Liu, B., Chiu, S.Y., Costantini, F. & Hsu, W. Development of a unique system for spatiotemporal and lineage-specific gene expression in mice. *Proc Natl Acad Sci U S A* **102**, 8615-20 (2005).

Zambrowicz, B.P. *et al*. Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci U S A* **94**, 3789-94 (1997).

Zeidner, N.S. *et al*. Sustained-release formulation of doxycycline hyclate for prophylaxis of tick bite infection in a murine model of Lyme borreliosis. *Antimicrob Agents Chemother* **48**, 2697-9 (2004).

Zhao, H.F., Boyd, J., Jolicoeur, N. & Shen, S.H. A coumermycin/novobiocin-regulated gene expression system. *Hum Gene Ther* **14**, 1619-29 (2003).

Zhu, P. *et al*. Silencing and un-silencing of tetracycline-controlled genes in neurons. *PLoS ONE* **2**, e533 (2007).

Zhu, Z., Ma, B., Homer, R.J., Zheng, T. & Elias, J.A. Use of the tetracycline-controlled transcriptional silencer (tTS) to eliminate transgene leak in inducible overexpression transgenic mice. *J Biol Chem* **276**, 25222-9 (2001).

Danksagung

Die vorliegende Arbeit wurde unter Anleitung von Frau Dr. Dagmar Wirth in der Arbeitsgruppe „Modellsysteme für Infektion und Immunität“ am Helmholtz Zentrum für Infektionsforschung in Braunschweig angefertigt. Mein besondere Dank gilt daher Dagmar für die Ermöglichung dieser Dissertation, für die ausgezeichnete Betreuung, Unterstützung, stete Bereitschaft zur Diskussion und nicht zuletzt für unermüdliches Korrekturlesen.

Ich danke Herrn Prof. Dr. Wehland und Herrn Prof. Dr. Steinert für die freundliche Übernahme des Referats bzw. Ko-Referats dieser Arbeit. Herrn Prof. Dr. Dübel danke ich für seine Bereitschaft zur Teilnahme an der Prüfungskommission.

Besonderer Dank gebührt Dr. Martin Hafner aus der Arbeitsgruppe „Experimentelle Immunologie“ für seine Anregungen und hilfreiche Unterstützung in Sachen mES Zellkultur und für die Organisation der Blastocysten-Injektionen, sowie Maria Ebel für deren technische Durchführung.

Dr. Hannes Klump von der Abteilung „Experimentelle Hämatologie“ der Medizinischen Hochschule Hannover unter der Leitung von Prof. Dr. Christopher Baum danke ich für die geduldige Einweisung in die ES-HC Differenzierung.

Maria Höxter möchte ich für ihren kompetenten Einsatz am FACS danken.

A very big thank you goes to the MSYS group formerly known as HIV/pig lab for their tolerance towards my little (?) quirks and for creating a work climate that left absolutely nothing to be desired. Special thanks go to Sara Bantner for providing the data obtained from ROSA26 targetings.

Ich danke ganz herzlich meiner Familie für ihre Unterstützung und Frank Hoyer, ohne dessen Hilfe ich niemals mit den Produkten aus dem Hause Microsoft klar gekommen wäre.

Lebenslauf

Persönliche Daten

Name:	Pamela Riemer
Geburtsdatum:	16.11.1979
Geburtsort:	Uelzen
Staatsangehörigkeit:	deutsch
Familienstand:	ledig

Schulausbildung

08/1986 – 06/1992	Grundschule und Orientierungsstufe Ebstorf
08/1992 – 06/1999	Lessing Gymnasium Uelzen Schulabschluss: Abitur

Studium

10/1999-03/2004	Studium der Biologie an der TU-Braunschweig
08/2003-03/2004	Diplomarbeit am Institut für Genetik der TU-Braunschweig, Arbeitsgruppe von Prof. Dr. Rüdiger Cerff, unter Anleitung von Dr. Marie-Françoise Liaud; Thema: „Lokalisierung der Glycerinaldehyd-3-phosphat-Dehydrogenase, Isoform GapC4, in der Kieselalge *Phaeodactylum tricornutum* durch „green fluorescent protein“ nach stabiler Transformation mit GFP-Reportergenkonstrukten

Berufstätigkeit

04/2004-07/2007	Wissenschaftliche Mitarbeiterin (Doktorandin) am Helmholtz Zentrum für Infektionsforschung in der Arbeitsgruppe „Modellsysteme für Infektion und Immunität“ von Dr. Dagmar Wirth

Publikationen

Wirth, D; Gama-Norton, L; Riemer, P; Sandhu, U; Schucht, R; Hauser, H (2007): „Road to precision: recombinase-based targeting technologies for genome engineering”, Curr. Opin. Biotechnol.; 26. September 2007, Epub ahead of print

“Science is like sex: sometimes something useful comes out, but that is not the reason we are doing it.”

Richard Feynman